UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION	1b RESTRICTIVE MARKINGS: N/A
2a SECURITY CLASSIFICATION AUTHORITY: N/A	DISTRIBUTION / AVAILABILITY OF REPORT: Distribution Unlimited
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): #7H432	5 MONITORING ORGANIZATION REPORT NUMBER(S): N/A
6a NAME OF PERFORMING ORGANIZATION: The University of Mississippi Medical Center; 6b OFFICE SYMBOL (If applicable): N/A	7a NAME OF MONITORING ORGANIZATION: Office of Naval Research
6c ADDRESS (City, State, and ZIP Code): Department of Biochemistry, 2500 North State Street, Jackson, MS 39216-4505	7b ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217-5000
8a NAME OF FUNDING / SPONSORING ORGANIZATION: Office of Naval Research; 8b OFFICE SYMBOL (If applicable): ONR	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-87-K-0387
8c ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217-5000	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO: 61153N; PROJECT NO: RR0416; TASK NO: 441c-028; WORK UNIT ACCESSION NO:

11 TITLE (Include Security Classification)
(u) Secretory Polypeptides Encoded in Balbiani Ring Genes

12 PERSONAL AUTHOR(S)
Case, Steven T.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Annual	FROM 89/04/15 TO 90/04/14	90/04/15	14

16 SUPPLEMENTARY NOTATION
N/A

17 COSATI CODES		
FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Eukaryotic genes, secretory proteins, tandemly repeated amino acid sequences, in vitro assembly/disassembly of polypeptide fibers, protein-protein interactions

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The goal of this project is to learn the structure and assembly characteristics of a family of secretory proteins (SP's). SP's are synthesized in salivary glands of aquatic larvae of the midge, Chironomus. Larvae construct underwater tubes for filter feeding and pupation using silk fibers composed of SP's. During the past year we made progress in two areas. (1) A 1-kb cDNA clone was sequenced that represents a portion of a 4.8-kb poly(A)+ RNA. Though it lacks sequences found in other SP mRNA's, its abundance in salivary glands is equal to other SP mRNA's. Antipeptide antibodies are being raised to identify the protein encoded by this mRNA. (2) Infrared and circular dichroism spectrometry have been used to study the structure of synthetic peptides representing the alternating "constant" and "subrepeat" domains of spI's. These 1000-kDal proteins appear to consist of alternating domains of α- and poly(Gly)II-helices.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	(u)
22a NAME OF RESPONSIBLE INDIVIDUAL: Dr. M. Marron	22b TELEPHONE (Include Area Code): (202)696-4760; 22c OFFICE SYMBOL: ONR

DD Form 1473, JUN 86 — Previous editions are obsolete — SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

Distribution List for Annual and Final Reports

1. Put a cover page (Form DD 1473) on your report and attach a copy of the distribution list. Mail one copy of the report to each person on the contractor subset list attached on which your name appears. The other subset list is for your information only. Please don't forget to attach this distribution list to your report - otherwise the folks below think they have mistakenly received the copy meant for the Molecular Biology Program and forward it to us.

2. Mail two copies to (include a DTIC Form 50 with these two copies too)
   Administrator
   Defense Technical Information Center
   Building 5, Cameron Station
   Alexandria, VA 22314

3. Mail one copy to each of the following:

(a) Dr. Michael Marron
ONR Code 1141
Molecular Biology Program
800 N. Quincy Street
Arlington, VA 22217-5000

(b) Administrative Contracting Officer
ONR Resident Representative
(address varies - see copy of your grant)

(c) Director,
Applied Research Directorate
ONR Code 12
800 N. Quincy Street
Arlington, VA 22217-5000

(d) Director
Office of Naval Technology
Code 22
800 N. Quincy Street
Arlington, VA 22217-5000

(e) Director
Chemical and Biological Sci Div
Army Research Office
P. 0. Box 12211
Research Triangle Park, NC 27709

(f) Life Sciences Directorate
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC 20332

(g) Director
Naval Research Laboratory
Technical Information Div, Code 2627
Washington, DC 20375

AMZEL, L. Mario
Department of Biophysics
Johns Hopkins School of Medicine
725 North Wolfe Street
Baltimore, MD 21205

ANDERSEN, Niels H.
Department of Chemistry
University of Washington
Seattle, WA 98195

ARNOLD, Frances H.
Dept of Chemical Engineering
California Institute of Technology
Pasadena, CA 91125

AUGUST, J. Thomas
Department of Pharmacology
Johns Hopkins Medical School
725 North Wolfe Street
Baltimore, MD 21205

BEVERIDGE, David L
Department of Chemistry
Wesleyan University
Hall-Altwater Laboratories
Middletown, CT 06457

BRAMSON, H. Neal
Department of Biochemistry
Univ of Rochester Medical Center
601 Elmwood Avenue
Rochester, NY 14642

BRUICE, Thomas C.
Department of Chemistry
University of California-Santa Barbara
Santa Barbara, CA 93106

CASE, Steven T.
Department of Biochemistry
Univ of Mississippi Medical Center
2500 North State Street
Jackson, MS 39216-4505

CHANG, Eddie L.
Bio/Molecular Engineering
Naval Research Laboratory
Code 6190
Washington, D.C. 20375-5000

CHRISTIANSON, David W.
Department of Chemistry
University of Pennsylvania
231 South 34th Street
Philadelphia, PA 19104-6323

CORDINGLEY, John S.
Department of Molecular Biology
University of Wyoming
Box 3944 University Station
Laramie, WY 82071

DeGRADO, William F.
E. I. du Pont de Nemours & Co
Central R & D, Experimental Station
P. O. Box 80328
Wilmington, DE 19880-0328

EVANS, David R.
Department of Biochemistry
Wayne State Univ School of Medicine
540 E. Canfield Street
Detroit, Michigan 48201

FEIGON, Juli F.
Department of Chem & Biochemistry
UCLA
405 Hilgard Avenue
Los Angeles, CA 900024-1569

FICHT, Allison R.
Dept of Med Biochem & Genetics
Texas A&M University
College Station, TX 77843

FRAUENFELDER, Hans
Department of Physics
University of Illinois
Urbana, IL 61801

GABER, Bruce
Naval Research Laboratory
Bio/Molecular Engineering Branch
Code 6190
Washington, DC 20375

GETZOFF, Elizabeth D.
Scripps Clinic & Research Foundation
Department of Molecular Biology
10666 North Torrey Pines Road
La Jolla, CA 92037

GOODMAN, Eugene M.
Biomedical Research Institute
University of Wisconsin
P. O. Box 2000
Kenosha, WI 53141

HO, Pui Shing
Department of Biochemistry and Biophysics
Oregon State University
Corvallis, OR 97331

HOGAN, Michael E.
Baylor Center for Biotechnology
4000 Research Forest Drive
The Woodlands, TX 77381

HONIG, Barry
Columbia University
Dept of Biochem and Molec Biophys
630 West 168th St.
New York, NY 10032

HOPKINS, Paul B.
Department of Chemistry
University of Washington
Seattle, WA 98195

KAHNE, Daniel
Department of Chemistry
Princeton University
Princeton, NJ 08544

KEMP, Robert G.
Chicago Medical School
Dept of Biological Chemistry
3333 Green Bay Rd.
North Chicago, IL 60064

KHORANA, Gobind H.
Department of Biology
MIT
77 Massachusetts Ave.
Cambridge, MA 02139

KIM, Sangtae
Chemical Engineering
University of Wisconsin
1415 Johnson Drive
Madison, WI 53706

LANSBURY, Peter T.
Department of Chemistry
MIT
Cambridge, MA 02139

LAURSEN, Richard A.
Chemistry Department
Boston University
590 Commonwealth Avenue
Boston, MA 02215

LENZ, Robert W.
Chemical Engineering Department
University of Massachusetts
Amherst, MA 01003

LEWIS, Randolf V.
Molecular Biology Department
University of Wyoming
University Station Box 3944
Laramie, WY 82071

LINDSAY, Stuart M.
Department of Physics
Arizona State University
Temp, AZ 85278

LOEB, George I.
David W. Taylor Research Center
Code 2841
Annapolis, MD 21402-5067

MASILAMANI, Divakar
Biotechnology Department
Allied-Signal Inc.
P. O. Box 1021R
Morristown, NJ 07960

McCONNELL, Harden M.
Stanford University
Department of Chemistry
Stanford, CA 94305

McELROY, Willam D.
Department of Chemistry
University of California - San Diego
La Jolla, CA 92093-0601

MERTES, Kristin Bowman
University of Kansas
Dept of Chemistry
Lawrence, Kansas 66045

NAGUMO, Mark
Bio/Molecular Engineering Branch
Naval Research Laboratory
Code 6190
Washington, DC 20375-5000

OLIVERA, Baldomero M.
Department of Biology
University of Utah
Salt Lake City, UT 84112

PABO, Carl O.
Department of Biophysics
Johns Hopkins University
School of Medicine
Baltimore, MD 21205

PRENDERGAST, Franklyn G.
Dept of Biochemistry & Molec Biol
Mayo Foundation
200 First St. SW
Rochester, MN 55905

PUGH, Jr., Edward N.
Deaprtment of Psychology
University of Pennsylvania
3815 Walnut Street
Philadelphia, PA 19104-6196

RACKOVSKY, Shalom R.
Department of Biophysics
University of Rochester
School of Medicine and Dentistry
Rochester, NY 14642

RAJAN, K. S.
Illinois Institute of Technology
Research Institute
10 W. 35th St.
Chicago, IL 60616

REINISCH, Lou
Laser Biophysics Center
Uniformed Services University
4301 Jones Bridge Road
Bethesda, MD 20814

RICH, Alexander
MIT Department of Biology
Cambridge, MA 02139

RICHARDS, J. H.
California Institute of Technology
Division of Chemistry and Chemical Engineering
Pasadena, CA 91125

ROTHSCHILD, Kenneth J.
Department of Physics
Boston University
590 Commonwealth Avenue
Boston, MA 02215

SCHULTZ, Peter G.
Department of Chemistry
University of California-Berkeley
Bekeley, CA 94720

SEEMAN, Nadrian
Department of Chemistry
New York University
New York, NY 10003

SELSTED, Michael E.
UCLA
Dept of Medicine
37-055 CHS
Los Angeles, CA 90024

SIGMAN, David S.
UCLA School of Medicine
Dept of Biological Chemistry
Los Angeles, CA 90024
SIKES, Steven C.
Department of Biological Sciences
University of South Alabama
Mobile, AL 36688

SINSKEY, Anthony J.
Laboratory of Applied Microbiology
MIT Department of Biology
Cambridge, MA 02139

STEWART, James M.
Department of Chemistry
University of Maryland
College Park, MD 20742

STEWART, John M.
Department of Biochemistry
University of Colorado
Health Science Center
Denver, CO 80262

TURNER, Douglas H.
Department of Chemistry
University of Rochester
Rochester, NY 14627

URRY, Dan W.
Laboratory of Molecular Biophysics
University of Alabama
P. O. Box 311
Birmingham, AL 35294

WAITE, J. Herbert
College of Marine Studies
University of Deleware
Lewes, DE 19958

WARD, Keith B.
Naval Research Laboratory
Code 6030
Washington, DC 20375

WARSHEL, Arieh
Department of Chemistry
University of Southern California
University Park
Los Angeles, CA 90089-0482

WATT, Gerald D.
Dept of Chemistry & Biochemistry
University of Colorado
Campus Box 215
Boulder, CO 80309-0215

ANNUAL REPORT
ONR Contract #N00014-87-K-0387
89 APR 15 through 90 APR 14

SECRETORY POLYPEPTIDES ENCODED BY BALBIANI RING GENES

Steven T. Case
Department of Biochemistry
The University of Mississippi Medical Center
Jackson, MS 39216-4505

Project Summary

The aim of this project is to learn about the structure, developmentally regulated synthesis and assembly of a family of secretory proteins (SPs) into an insoluble polymer of silk-like threads. SPs are exclusively synthesized in salivary glands of aquatic larvae of the Dipteran, Chironomus. All SPs studied to date are composed of tandemly repeated amino acid sequences. Recombinant cDNA probes are used to map SP-coding genes on polytene chromosomes, identify their mRNAs on Northern blots and derive the amino acid sequence of their encoded protein. cDNA probes and anti-SP antibodies are used to study the level at which SP gene regulation occurs during larval development and under conditions of galactose-induced alterations in gene expression. SP assembly in vitro is being studied by a combination of physical, electron microscopic and biochemical methods. We hope to learn which SPs interact with each other and what is the chemical nature of these interactions. We eventually plan to determine the spatial distribution of SPs within assembled complexes by making three-dimensional tomographic reconstructions from immunoelectron micrographs. This experimental system provides a unique opportunity to study how naturally occurring soluble proteins can assemble into an insoluble fiber that functions in an aqueous environment.

Results from the Prior Year

cDNA clones for additional SPs. Manuscripts describing the identification, partial structure and expression of genes for sp140 (Dignam et al., 1989) and sp185 (Dignam and Case, 1990) are attached. During the past year, another partial cDNA clone (F4a) has been sequenced. While the 1-kb sequence lacks landmarks of other SP-coding cDNAs (tandemly repeated sequences, conserved Cys residues, + Pro - motifs), it hybridizes to a 4.8-kb poly(A)+ RNA that in salivary glands appears to be as abundant as SP-coding mRNAs . To identify the protein encoded by this mRNA we are raising antibodies against a 16-residue synthetic peptide whose sequence was derived from the cDNA. We are experimenting with an approach developed by Tamm and co-workers whereby peptide conjugation to a carrier protein is circumvented by initiating peptide synthesis on a branched-chain, octavalent core of Lys residues.

In vitro assembly/disassembly of SPs. A manuscript describing the structure and in vitro disassembly/reassembly of SPs into macromolecular complexes has been published (Wellman and Case, 1989). We demonstrated that purified spIs (the 1000-kDa SPs) could assemble into complexes with both similar morphology, using electron microscopy, and a similar dichroic spectrum to that of unfractionated native complexes. We estimated from quantitative dichroism measurements that spIs were approximately 15% α-helix, 28-30% β-sheet, 26-28% β-turn and 25% other secondary structure. During the past year we synthesized and purified peptides corresponding to the alternating CONSTANT (C) and SUBREPEAT (SR) domains of spIs. C and SR peptides

were subjected to circular dichroism and infrared spectrometry. Data for the C peptide are consistent with predictions of its secondary structure; it consists primarily of α-helix. The SR peptide, however, does not consist of ß-structure. Instead it appears to form a left-handed, 3_{10}/poly(Gly)II-type helix. Finally, the Amide A frequencies suggest that both C and SR peptides may form sheets of parallel helices rather than multistranded, supercoiled structures.

Interspecific comparisons of SPs. We initiated a comparison of SP structure among related but different species of Chironomus. Hybridization and antibody probes have been exchanged with: Professor J.-E. Edstrom's lab (Lund, Sweden) which studies C. pallidivittatus; Professor I.I. Kiknadze's lab (Novosibirsk, U.S.S.R.) which studies C. thummi. Aside from identifying homolgous SP proteins in each species, we hoped that species-specific differences in the synthesis of SPs might provide clues to the role of each protein in silk fiber assembly. While working in Kiknadze's lab as a U.S.-U.S.S.R. National Academies of Sciences Exchange Scientist, I made an interesting observation. Silk fibers spun by larvae from each species are distinguishable microscopically. The morphological differences that were observed correlate with the presence of a cell-specific SP that is not synthesized by C. tentans larvae.

Plans for Next Year

SP-coding genes. We will continue to identify and map additional SP-coding genes using recombinant DNA procedures which have been successful to date. Our first priority will be to determine if the F4a antipeptide antibody reacts with an SP on Western blots. If it does, we will continue to study the structure and expression of this mRNA. Our next priority will be to construct a cDNA expression library from C. thummi salivary gland RNA. We want to clone the cDNA for the cell-specific SP mentioned above.

Structural organization of SP complexes. We will continue to study SP-SP interactions focusing on the C and SR domains of spIs. Biophysical studies of synthetic C and SR peptides will continue. Their assembly characteristics will be compared to native spIs purified by glycerol gradient centrifugation. We hope to identify one or more sites within these domains that are responsible for in vitro assembly. For example, preliminary data suggest that C peptides can dimerize by formation of a disulfide bond. Peptide mapping will be used to determine whether this disulfide bond involves a specific or random pair of Cys residues. Similar studies will be extended to spI complexes formed in vitro and SP complexes formed in vivo.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 264, No. 16, Issue of June 5, pp. 9444-9452, 1989
Printed in U.S.A.

Identification of a Developmentally Regulated Gene for a 140-kDa Secretory Protein in Salivary Glands of *Chironomus tentans* Larvae*

(Received for publication, December 19, 1988)

Susan S. Dignam‡, Lily Yang, Markus Lezzi§, and Steven T. Case¶

From the Department of Biochemistry, University of Mississippi Medical Center, Jackson, Mississippi 39216-4505 and §Institute for Cell Biology, Swiss Federal Institute of Technology, Honggerberg, CH-8093 Zurich, Switzerland

Secretory proteins are synthesized in salivary glands of the insect, *Chironomus tentans*, and assemble *in vivo* into silk-like threads which aquatic larvae use to construct tubes for filter feeding and pupation. Thus far, all known secretory protein genes contain repetitious protein-coding sequences and are located in cytological structures known as Balbiani rings, giant puffs found on polytene secretory cell chromosomes. In this paper we describe the identification of another secretory protein gene which is comprised of repeated sequences; however, this gene is not located in a Balbiani ring. Two partial cDNA clones from a 3.6-kilobase pair poly(A)$^+$ RNA were sequenced and found to contain two open reading frames for protein synthesis. Antibodies were raised against synthetic oligopeptides whose sequences were derived from these two open reading frames. An immunoaffinity-purified antibody for one of these peptides bound specifically to a 140-kDa secretory protein (sp140). The cDNA sequences contain tandem repeats of 42 base pairs which encode a repeat of 14 amino acids with a composition and oligopeptide sequence similar to other secretory proteins. The *C. tentans* genome contains about 70 copies of this 42-base pair repeat organized as a contiguous block of 3 kilobase pairs or less. The sp140 gene was mapped by *in situ* hybridization to polytene chromosome band I-17-B. Developmental studies of protein accumulation, steady-state levels of mRNA, and relative transcription rate suggested that the sp140 gene is developmentally regulated so that maximal expression is achieved during the prepupal stages of the fourth larval instar. Based upon these results we proposed that sp140 gene belongs to a prepupal class of secretory protein genes. While the sp140 gene shares structural and expression characteristics with other secretory protein genes, its unique chromosomal location shows that this multigene family is not restricted to Balbiani rings.

The organization and chromosomal distribution of eukaryotic multigene families is paradoxical. Some gene families (1–9) are dispersed throughout the genome, while others (10–16) are organized as gene clusters on one or more chromosomes. In some instances, the organization and distribution of gene clusters may be important for coordinate regulation of the expression of its constituent genes. For example, the spatial distribution of genes in the globin cluster reflects their temporal pattern of expression during development (12). Certain histone gene clusters represent different expression classes of genes (10). In other instances, there is no apparent pattern of expression for genes within a cluster.

The chromosomal distribution of the secretory protein multigene family in *Chironomus* contributes to this paradox. For example, the spI[1] class of genes (17–19 and references therein) consists of four members which encode structurally similar secretory proteins (spIa, spIb, spIc, and spId). Each protein has a molecular mass of about 1000 kDa and is largely composed of complex core repeats which can be divided into two regions. The constant (C) region is 35–45 amino acids in length and contains 4 cysteine, 1 methionine, and 1 phenylalanine residues which are invariant. The subrepeat (SR) region contains four–six direct repeats of 6–12 amino acid sequences which contain a characteristic tripeptide motif: a positively charged residue (lysine or arginine) followed by proline followed by a negatively charged residue (glutamic acid, aspartic acid, or phosphoserine), henceforth designated as ⊕ Pro ⊖. These alternating C and SR regions are thought to be important sites for protein-protein interactions which take place when secretory proteins assemble into silk-like threads (18, 19). Larvae spin these threads to construct underwater tubes for filter-feeding and pupation (20, 21). All spI genes are located in cytological structures known as Balbiani rings (BRs), and their distribution on polytene chromosomes reflects their differential expression in response to galactose (22–26): BR1 on chromosome IV contains the spIa gene whose expression is generally unaffected by galactose; BR2 on chromosome IV contains the spIb and spId genes both of which are repressed by galactose; BR6 on chromosome III contains spIc, the only gene which is induced by galactose.

BR1, like BR2, also contains a cluster of secretory protein genes. In addition to the spIa gene, it contains the gene for

* This research was supported by Office of Naval Research Contract N00014-87-K-0387 (to S. T. C.) and a grant from the Swiss National Science Foundation (to M. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J04949.

‡ Present address: Dept. of Biochemistry, Medical College of Ohio, C.S. 10008, Toledo, OH 43699.

¶ To whom all correspondence should be addressed: Dept. of Biochemistry, the University of Mississippi Medical Center, 2500 North State St., Jackson, MS 39216-4505. Tel.: 601-984-1518.

[1] The abbreviations used are: spI, a 1000-kDa secretory protein; sp"x", other secretory proteins with an apparent M_r equal to "x"; BR, Balbiani ring; sp, secretory protein; C region, constant region of an spI core repeat; SR region, subrepeat region of an spI core repeat; ORF, open reading frame; bp, base pair; kb, kilobase pairs.

sp195 (27, 28). This secretory protein is largely composed of simple direct repeats of 25 amino acids which include one copy of the tripeptide motif found in sp1 SR regions. In direct contrast to sp1 genes, which can be expressed throughout all stages of larval development, the expression of sp195 is limited to the prepupal stages (29). The developmentally regulated expression of sp195 refutes the notion that the cluster of secretory protein genes in BR1 represents a single expression class of genes.

We have continued to identify, study the expression of, and map secretory protein genes in *Chironomus* with the following questions foremost in our minds. Are all secretory protein genes clustered in BRs? Do all secretory proteins contain repeated amino acid sequences? Can any pattern of chromosomal distribution and expression of members of this multigene family be discerned? We report here the identification of a gene for sp140 which is largely composed of repeated sequences and whose expression is developmentally regulated. Surprisingly, however, this secretory protein gene is not located in a BR.

MATERIALS AND METHODS

Staging of Larvae—Individual *Chironomus tentans* larvae were staged based upon the morphology and orientation of imaginal disks (30). Salivary glands were removed from staged larvae, placed in 70% ethanol on ice for 1 h, and then stored at −20 °C for up to several weeks while all other samples were collected. Developmental studies of mRNA and protein levels were done with salivary gland extracts made from pools of staged larvae taken from individual egg strings.

Purification of Nucleic Acids—Routine procedures described without specific literature citations followed protocols similar to those described by Maniatis *et al.* (31). Total RNA was extracted (32) from salivary glands dissected from fourth instar larvae. Poly(A)$^+$ RNA was obtained by chromatography of salivary gland RNA over oligo(dT)-cellulose. Genomic DNA was extracted (33) from crude nuclei of *C. tentans* tissue culture cells (34) and purified by density gradient centrifugation in CsCl. Recombinant plasmid DNAs were purified from bacterial lysates in CsCl gradients containing ethidium bromide. Oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer, deblocked, and purified on 20% polyacrylamide gels containing 7 M urea. Oligonucleotide probes used in this study were: C5.1 (5′ CCTTTGTTAGGAGCTTTGCC-3′, the reverse complement of nucleotides 17–36 in Fig. 3); C5.2 (5′-GGTAACA-ACGCAAAGAGTTGC-3′, nucleotides 113–133 in Fig. 3).

Radiolabeling and Hybridization of Nucleic Acid Probes—Oligonucleotides were end-labeled with [γ-^{32}P]ATP and polynucleotide kinase. Recombinant plasmid DNAs were labeled by nick translation using DNA polymerase I and [α-^{32}P]dATP (for blot hybridization) or [α-^{35}S]dCTP (for *in situ* hybridization). Northern blots (35) of salivary gland RNA were made from denaturing 0.75% agarose gels containing methyl mercury hydroxide (36). RNA markers were obtained from Bethesda Research Laboratories (Catalogue 5620SA) and run in parallel lanes of each gel. Southern blots (37) of restricted genomic and plasmid DNAs were made from agarose gels containing *Hin*dIII fragments of bacteriophage λ DNA (38) or *Hae*III fragments of pBR322 DNA (39) as markers. Dot-blots of RNA (26) and DNA (40) were made as described. All hybridizations were done in SET (1 × SET is 0.15 M NaCl, 2 mM EDTA, 30 mM Tris-HCl, pH 8.0) solutions containing 0.1% sodium dodecyl sulfate, 0.1% potassium pyrophosphate, and 500 μg/ml heparin. Oligonucleotide probes were hybridized 24–48 h in 4 × SET at 50 °C, rinsed in 2 × SET four times for 15 min at room temperature and rinsed once for 5 min at 50 °C. Plasmid DNA probes were hybridized in 4 × SET at 65 °C, rinsed in 4 × SET at room temperature twice for 30 min and rinsed in 1 × SET for 1 h at 65 °C. ^{35}S-Labeled cDNA plasmids were hybridized *in situ* to squashed preparations (41) of salivary gland polytene chromosomes and detected by autoradiography and staining (42).

Quantification of Hybridization Data—A series of timed autoradiographic exposures were made from Southern, Northern, and dot blots that were hybridized with radioactive probes. Two dimensional autoradiographic images were scanned and digitized using a Bio-Rad model 620 video densitometer connected to an IBM PC-AT computer. Data collection, image processing, and conversion of image intensities to volumetric densitometric units were accomplished using Bio-Rad's *two dimensional Analyst software*. When specified, blots also contained, as a reference, a serial dilution of known quantities of plasmid or phage DNA with or without cDNA inserts. The reference images ensured that hybridization probes were in excess and autoradiographic images were within the linear exposure range of the film.

DNA Sequencing—The construction of plasmids containing cDNA that was randomly primed from fractions of salivary gland RNA has been described (27). cDNA inserts, flanking homopolymeric tails, and adjacent sequences from pBR322 were subcloned in both orientations into the *Sma*I site of the replicative form of bacteriophage M13mp18 DNA (43) by blunt-end ligation. DNA sequencing (44) was done on full length inserts and deletion constructs made with exonuclease III (45). Sequencing primers contained either M13 or cDNA sequences that were obtained during the course of these experiments. ^{35}S-Labeled sequencing products were separated on 6–8% polyacrylamide gradient gels (46). All reactions were done at least twice, and the complete sequence of both strands of each insert was determined independently. DNA sequences were compiled and analyzed using Pustell's (47) programs which were obtained from International Biotechnologies, Inc.

Oligopeptide Synthesis and Affinity Purification of Antipeptide Antibodies—Oligopeptides were synthesized using standard methods for solid phase chemical peptide synthesis (48, 49). The sequence selected for the +2-peptide was NH$_2$-Lys-Pro-Glu-Gly-Lys-Ala-Pro-Asn-Lys-Gly-Lys-Asn-Gly-Ser-COOH (encoded by nucleotides 50–91 in Fig. 3); the sequence selected for the +3-peptide was NH$_2$-Lys-Ala-Arg-Met-Glu-Ala-Asn-Gln-Lys-Ala-Lys-Leu-Gln-Val-COOH (encoded by nucleotides 75–116 in Fig. 3). Peptides were cleaved from the support resin, covalently cross linked to bovine serum albumin with glutaraldehyde and injected into rabbits to obtain polyclonal antipeptide antisera (28). Antipeptide antibodies were immunoaffinity-purified (28) by sequentially chromatographing antisera over columns of Affi-Gel 10 (Bio-Rad) which contained either BSA or the peptide that was used as the immunogen.

Gel Electrophoresis and Immunoblotting—Secretory proteins were extracted from the lumen of salivary glands fixed in 70% ethanol (28, 50) in 6 M guanidine hydrochloride, reduced, alkylated, and fractionated by gel electrophoresis on 3–20% concave exponential gradients of polyacrylamide (25). The sample buffer, however, contained Pyronin Y as the tracking dye which was run to the bottom of the gel. Western blots were made by electrophoretic transfer of proteins to nitrocellulose (51) and stained with Ponceau S (52). Blots were photographed, destained, and reacted with affinity purified antipeptide antibodies as described (28). Bound primary antipeptide antibodies were detected by their reaction with secondary goat anti-rabbit antibodies coupled to alkaline phosphatase (53).

Run-on Nuclear Transcription Assays—For each nuclear transcription assay, nuclei were isolated (54) from 10 salivary glands obtained from staged larvae. Nuclei were incubated 45 min at 29 °C in 2 μl of buffer (54) containing 50 μM [α-^{32}P]UTP. ^{32}P-Labeled nuclear RNA was extracted (32) and hybridized for 72 h at 65 °C in 4 × SET (as described above) to dot blots of recombinant M13 phage DNA containing either the transcribed or nontranscribed strand of cDNA. The posthybridization rinses included a 1-h incubation with pancreatic ribonuclease. Radioactive hybrids were detected by autoradiography and compared by two-dimensional densitometry of the autoradiographic images, as described above.

RESULTS AND DISCUSSION

Isolation of Partial cDNA Clones from a Salivary Gland 3.6-kb Poly(A)$^+$ RNA—We previously described the synthesis of randomly primed cDNA from salivary gland RNA and its cloning by annealing to the *Pst*I site of pBR322 using (dA)-(dT) homopolymeric tails (27). Several clones were selected for analysis based upon the observation that they exhibited relatively intense autoradiographic signals after colony hybridization with ^{32}P-labeled salivary gland poly(A)$^+$ RNA. Recombinant plasmid DNA from one of these clones was purified, radiolabeled by nick translation, and hybridized to a Northern blot of salivary gland RNA. This plasmid, designated pCt140.1, hybridized to a 3.6-kb poly(A)$^+$ RNA (Fig. 1*A*). Based upon our prior experience with secretory protein genes (27), the autoradiographic intensity of this band indicated that the transcript was abundant and/or contained internally repetitive sequences.

pCt140.1 was partially characterized by restriction mapping, and it was estimated that the insert contained about 300 bp of DNA. Since the plasmid's *Pst*I site was destroyed by our cloning strategy, the cDNA and adjacent sequences from pBR322 were subcloned as two separate *Hae*III fragments by blunt-end ligation into the *Sma*I site of the replicative form of bacteriophage M13mp18. Fragments cloned in one or the other orientation yielded M13 phages that provided strand-specific probes which, when hybridized to Northern blots (data not shown), enabled us to deduce the transcriptional polarity of the cDNA (Fig. 2).

The nucleotide sequence of the complete insert (Fig. 3) was obtained from dideoxyribonucleotide-terminated extensions of a universal M13 sequencing primer which was annealed to a series of deletion constructs that were made with exonuclease III (Fig. 2). The cDNA sequence was also used to design the synthesis of two strand-specific oligonucleotide hybridization probes, C5.1 and C5.2 (see "Materials and Methods"). These probes were used to (*a*) confirm the transcriptional polarity of the cDNA by hybridization to Northern blots (Fig. 1*B*), (*b*) prime sequencing reactions on double-stranded pCt140.1 DNA to confirm the sequence at the junction of the subcloned *Hae*III fragments (see Fig. 2), and (*c*) rescreen the partial cDNA library for clones with related sequences.

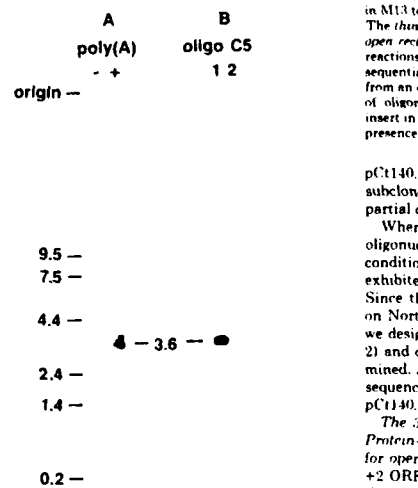

FIG. 1. **Hybridization specificity of cDNA and oligonucleotide probes to salivary gland RNA.** Northern blots were made from denaturing 0.75% agarose gels containing (*A*) 1 μg of poly(A)$^+$ and poly(A)$^-$ RNA and (*B*) 5 μg of total salivary gland RNA. Autoradiograms show the hybridization patterns obtained after the blots were hybridized with ^{32}P-labeled (*A*) pCt140.1 or (*B*) oligonucleotides C5.1 and C5.2 (see "Materials and Methods"). *Numbers* to the *left* indicate the size (in kilobase pairs) of RNA markers that were run in parallel lanes and detected by staining the gel with ethidium bromide.

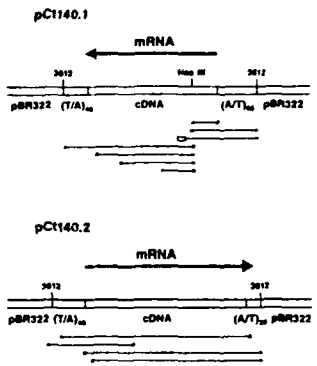

FIG. 2. **The structure, transcriptional polarity, and strategy used for sequencing recombinant cDNA plasmids.** The diagrams of two different plasmids (pCt140.1 and pCt140.2) show the relative size and orientation of the cDNA and flanking homopolymeric tails that were inserted into the *Pst*I cleavage site (nucleotide 3612) of pBR322. These data were obtained by a combination of DNA sequencing and the ability to discriminate between single-stranded cDNA subclones in M13mp18 (made from *Hae*III fragments generated by cleavage at nucleotides 3554 and 3654 in pBR322) by their differential hybridization with oligonucleotides from either side of the *Pst*I site. The *thick arrow* above each diagram shows the direction of transcription. This was deduced by the ability of cDNA subclones in M13 to hybridize with oligonucleotide probes C5.1 or C5.2 (Fig. 1). The *thin arrows* below each diagram show the origin (*filled circles* or *open rectangle*), direction, and extent of sequencing reactions. Most reactions were done on single-stranded M13 templates containing sequential deletions made with exonuclease III. The *arrow* originating from an *open rectangle* represents a reaction primed by hybridization of oligonucleotide C5.2 to the full-length double-stranded cDNA insert in pCt140.1. This reaction was used to overlap and confirm the presence of the *Hae*III site (see Fig. 3).

When the cDNA library was rescreened with ^{32}P-labeled oligonucleotide C5.1 under relatively stringent hybridization conditions (T_m, −6 °C) another clone was selected which exhibited greater autoradiographic intensity than pCt140.1. Since this plasmid also hybridized to 3.6-kb poly(A)$^+$ RNA on Northern blots of salivary gland RNA (data not shown), we designated it pCt140.2. The transcriptional polarity (Fig. 2) and cDNA sequence (Fig. 3) of pCt140.2 were also determined. A comparison (presented below) of the 193-bp cDNA sequence of pCt140.1 and the 243-bp cDNA sequence of pCt140.2 revealed that they were extremely similar.

The 3.6-kb Transcript Is mRNA for a 140-kDa Secretory Protein—When the coding strand of each cDNA was searched for open reading frames (ORFs), two ORFs were found: the +2 ORF in pCt140.1 is similar to the +1 ORF of pCt140.2; the +3 ORF in pCt140.1 is similar to the +2 ORF in pCt140.2 (Fig. 3). To determine which of these ORFs were utilized for protein synthesis *in vivo*, antibodies were raised against synthetic oligopeptides whose sequences correspond to a portion of each ORF in pCt140.1. Each peptide was used to immunize rabbits. Rabbit antisera were immunoaffinity purified (28) to yield highly specific, anti-(+2-peptide) and anti-(+3-peptide) antibodies.

To determine if either of the peptide-reactive antibodies

---

^{1} Our most recent results suggest that previously used values (28, 29) for the apparent M_r of secretory proteins sp115 through sp180 were underestimated. In this manuscript, values will be used which reflect more accurately the relative electrophoretic mobilities of secretory proteins and reference proteins on polyacrylamide gels containing sodium dodecyl sulfate. To facilitate comparison to our previous work and without confusion, we point out here that sp195 replaced sp190, sp185 replaced sp170, sp165 replaced sp150, sp140 replaced sp130, and sp125 replaced sp115.

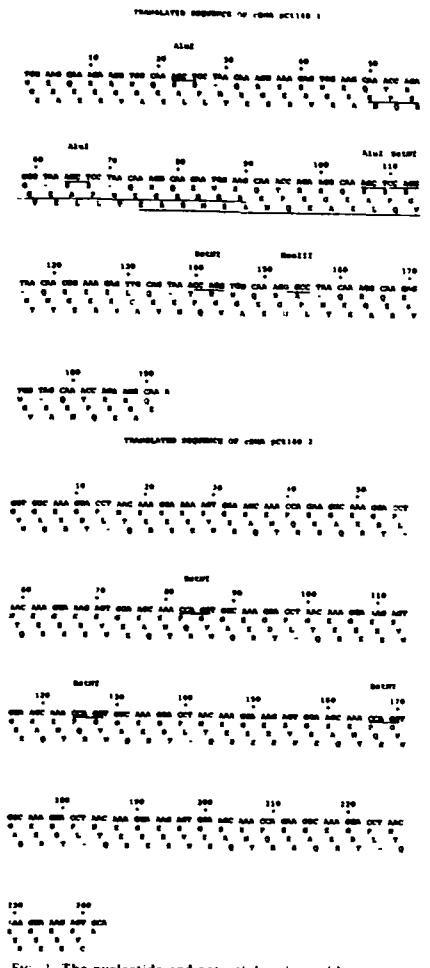

FIG. 3. **The nucleotide and potential amino acid sequences of cDNAs in pCt140.1 and pCt140.2.** The nucleotide sequence of the mRNA strand (Fig. 2) of each cDNA is presented as the *top line* of each panel and *numbered*. Nucleotides which comprise the *recognition sequences of AluI, BstNI, and HaeIII were underlined*. These sites were located by computer analysis of the cDNA sequences. Potential amino acid sequences derived from translating reading frames beginning with the first, second, and third nucleotide of each cDNA sequence are displayed using the conventional single letter code. *Dashes* in the amino acid sequences represent stop codons. *Underlined* amino acid sequences in pCt140.1 were selected for oligopeptide synthesis.

bound to salivary gland protein, aliquots of a denatured, reduced, and alkylated extract of secretory proteins were fractionated by polyacrylamide gel electrophoresis (25). Proteins were transferred electrophoretically to sheets of nitrocellulose and stained with Ponceau S (Fig. 4*A*). A spectrum of secretory proteins was visible including those with an apparent molecular mass of 35–1000 kDa. When affinity-purified anti-(+2 peptide) antibody was incubated with a destained lane of this blot, the antibody bound selectively to a single protein with an apparent molecular mass of 140 kDa (Fig. 4*B*). If, however, this antibody was preincubated with 50 μM +2-peptide, the antibody reaction was abolished (Fig. 4*C*). Whereas affinity-purified anti-(+3-peptide) antibody reacted specifically with its cognate peptide, it failed to react with any salivary gland protein (Fig. 4*D*).

These results led us to conclude that pCt140.1 is a partial cDNA clone derived from a 3.6-kb poly(A)$^+$ mRNA for sp140. *Furthermore, since both cDNA clones hybridize the same size* RNA and have nearly identical sequences (see below) we assumed that the partial amino acid sequence of sp140 is represented by the +2 ORF in pCt140.1 and the +3 ORF of pCt140.2.

sp140 Is Partially Composed of Tandem Copies of Repeated Sequences—Further analysis of the cDNA sequences from pCt140.1 and pCt140.2 revealed that they were composed of nearly identical, tandem repeats of a 42-bp sequence (Fig. 5). For the purpose of comparison, the nucleotide and deduced amino acid sequences (based upon data presented in Fig. 4) were displayed as repeat units (Fig. 5). Three repeats in pCt140.2 (nucleotide positions 55–96, 97–138, and 139–180, rows 3, 4, and 5, respectively) were identical and considered to represent a consensus repeat sequence. The other full-

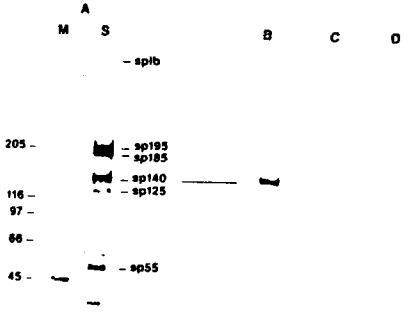

FIG. 4. **Immunological identification of the translational product encoded by cDNA sequences in pCt140.1.** *A*, aliquots of an extract of secretory (S) proteins were fractionated by slab gel electrophoresis in parallel with molecular weight markers (*M*). Proteins were electrophoretically blotted to a nitrocellulose membrane, stained with Ponceau S, and photographed. After destaining the filter, *various lanes were incubated with a* 1:20 *dilution of immunoaffinity*-purified rabbit antipeptide antibodies: *B*, anti-(+2 peptide) antibody; *C*, anti-(+2 peptide) antibody preincubated with 50 μM +2 peptide; *D*, anti-(+3 peptide) antibody. Primary antibodies were detected by their reaction with secondary goat anti-rabbit antibody conjugated with alkaline phosphatase. *Numbers* to the *left* correspond to the molecular mass (in kDa) of markers (from *top* to *bottom*: myosin, β-galactosidase, phosphorylase B, bovine serum albumin, and ovalbumin). Numbers between *A* and *B* correspond to the apparent molecular mass (in kDa) of secretory proteins (sp).

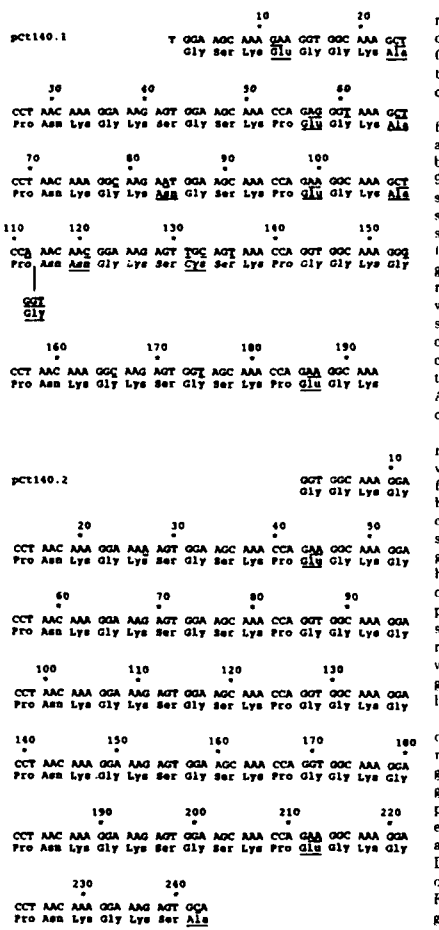

FIG. 5. **The organization of repeated sequences in the cDNA and polypeptide sequence of sp140.** Portions of the sp140 mRNA (see Fig. 2) and encoded polypeptide sequences (see Fig. 4) which are contained in cDNA plasmids pCt140.1 and pCt140.2 are displayed as tandem repeat units. *Underlined* nucleotides or amino acids deviate from the consensus sequence chosen from the three identical repeats in rows 3, 4, and 5 (nucleotides 55–180) of the pCt140.2 sequence. To achieve maximal alignment of the pCt140.1 sequence, we assumed that the GGT codon for glycine at nucleotides 113–115 in pCt140.1 represents an in-frame insertion of a codon.

length and half-repeats had between 93% (39/42 nucleotides between positions 13–54) to 95% (40/42 nucleotides between positions 181–222 and 20/21 nucleotides between positions 223–243) nucleotide sequence similarity with the consensus sequence. Whereas one base substitution resulted in a silent mutation (G to A at position 27) the others resulted in codon changes (G to C at position 242 changed glycine to alanine; a GT dinucleotide was changed to an AA dinucleotide at positions 44 plus 45 and 212 plus 213, both resulting in a codon change of glycine to glutamic acid).

The repeats in pCt140.1 were somewhat more divergent from the consensus repeats in pCt140.2 (Fig. 5). Both partial and complete repeat units displayed in rows 1, 2, 3, and 5 had between 84% (21/25 nucleotides between positions 1–25) and 90% (35/39 nucleotides between positions 155–193) sequence similarity with the pCt140.2 consensus repeat. The alignment shown for the repeat in row 4 had 86% (36/42 nucleotides) sequence similarity if one assumes that the GGT trinucleotide (positions 113–115) represents the insertion of an in-frame glycine codon. Six of the base substitutions were silent. The remaining 19 substitutions led to 10 codon changes, six of which were noteworthy because they involved dinucleotide substitutions which were sequestered at 2 residues: the GGT codon for Gly^{41} in the consensus repeat of pCt140.2 was changed to GAG or GAA codons for Glu^{41} (rows 2, 3, and 5); the GGA codon for Gly^{51} was changed to a GCT codon for Ala^{51} (rows 1, 2, and 3). Recall that pCt140.2 had two dinucleotide substitutions which also changed Gly^{41} to Glu^{41}.

In summary, cDNA in pCt140.1 and pCt140.2 contained nearly identical, tandem copies of a 42-bp consensus sequence which encoded a 14-residue amino acid sequence. The codons for residues 11 and 14 displayed more than half (17/30) of the base substitutions, and all but one of these were dinucleotide changes. In spite of this, there appeared to be selective pressure to maintain glycine or glutamic acid at residue 11 and glycine or alanine at residue 14. The sequence similarities between these two cDNAs suggest that they were either derived from two 3.6-kb mRNAs encoding rather similar proteins or they were primed from two locations within the same mRNA. If the latter idea is correct, then the divergent repeat in pCt140.1 (row 4) may prove particularly interesting with regard to understanding how repeats within the sp140 gene evolved. Such an analysis will await the isolation of full-length cDNA.

Genomic Organization and Location of the sp140 Gene—A dot-blot hybridization experiment was performed to determine how many copies of the 42-bp repeat were present in genomic *C. tentans* DNA. Dots containing a serial dilution of genomic DNA were spotted in parallel to a serial dilution of pCt140.1 DNA. The resulting blot was hybridized with a molar excess of ^{32}P-labeled oligonucleotide C5.1. Comparison of the autoradiographic intensities indicated that 1 ng of pCt140.1 DNA contains about the same number of copies of the C5.1 oligonucleotide as 1.4 μg of genomic DNA (data not shown). From these results it was calculated that the *C. tentans* genome contains about 70 copies of the C5.1 oligonucleotide sequence. This calculation was based upon the following assumptions: (*a*) the size of pCt140.1 is 4660 bp (including vector, homopolymeric tails, and cDNA); (*b*) pCt140.1 contains three copies of the C5.1 sequence which are sufficiently identical to hybridize the probe; and (*c*) the haploid DNA content of the *C. tentans* genome is 1×10^{11} daltons (55). If all 70 copies were arranged contiguously in one gene, this would predict a minimum gene length of 2.9 kb or more than 80% of the length of sp140 mRNA.

The copy number and overall organization of repeats was determined from Southern blots prepared from *C. tentans* genomic DNA cleaved with various restriction endonucleases. When blots from 0.7% agarose gels were hybridized with ^{32}P-labeled oligonucleotide C5.1, a simple pattern of autoradiographic bands was generally obtained (Fig. 6). Digestion by

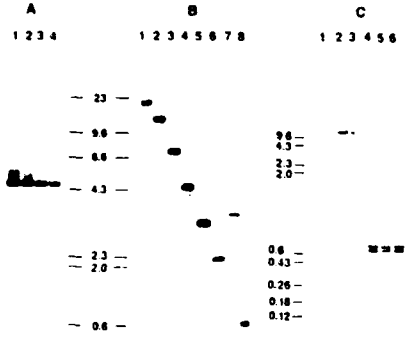

FIG. 6. **Genomic Southern blot analysis of the sp140 gene.** *A*, 5 ng (*1*), 2.5 ng (*2*), 1 ng (*3*), and 0.5 ng (*4*) of pCt140.1 DNA cleaved with *Eco*RI and run on a 0.7% agarose gel. *B*, 2.5-μg samples of *C. tentans* nuclear DNA cleaved with *Sal*I (*1*), *Bam*HI (*2*), *Eco*RI (*3*), *Mbo*I (*4*), *Taq*I (*5*), *Alu*I (*6*), *Bst*NI (*7*), or *Hae*III (*8*) and run on the same gel as *A*. *C*, 1-μg samples of *C. tentans* nuclear DNA partially cleaved with *Hae*III at 37 °C for 0 min (*1*), 5 min (*2*), 10 min (*3*), 20 min (*4*), 40 min (*5*), and 60 min (*6*) and separated on a 2% agarose gel. DNA was transferred from each gel to Nytran membrane and hybridized with ^{32}P-labeled oligonucleotide C5.1. Numbers along the sides of panels represent the position and size (in kilobase pairs) of DNA markers (*23–0.6* kb are *Hin*dIII fragments of λ DNA; *0.43–0.12* kb are *Hae*III fragments of pBR322).

enzymes such as *Sal*I, *Bam*HI, *Eco*RI, *Mbo*I, and *Taq*I resulted in single bands of between 3 and 20 kb. Densitometry of the autoradiographic images indicated that the intensity of bands obtained with 2.5 μg of *C. tentans* DNA was comparable to the intensity obtained from 2 ng of linearized pCt140.1 DNA (Fig. 4, *A* and *B*). Based upon assumptions discussed in the previous paragraph, the simplest interpretation of these results was that the genome contains 60–70 copies of the 42-bp repeat arranged in a single block as small as 3 kb.

Genomic Southern blots of *Alu*I-cleaved DNA also yielded a single band of about 2.2 kb with an intensity equivalent to about 20 copies of the 42 bp repeat (Fig. 6*B*). This observation combined with the apparent lack of lower molecular weight bands suggested that a portion of the repeat block was extensively cleaved by this enzyme. In fact, pCt140.1 contains three *Alu*I sites 42 bp apart, whereas pCt140.2 is devoid of these sites (Fig. 3). Similarly, *Bst*NI yields a 3.3-kb band of hybridization which we estimated contains 7–10 copies of the repeat (Fig. 6*B*). This result was consistent with the periodic distribution of *Bst*NI sites in both cDNA sequences (Fig. 3). Finally, *Hae*III apparently cleaved genomic DNA into an intense band of about 0.6 kb, and faint bands of 5 and 10 kb (Fig. 6*B*). In fact, densitometric tracings and higher concentration agarose gels revealed that the low molecular weight *Hae*III band seen on 0.7% agarose gels actually consisted of two equally intense bands of about 600 and 650 bp (Fig. 6*C*). We estimated that these *Hae*III bands each contain about 10–15 copies of the repeat. This result agrees with the occurrence of only a single *Hae*III site within either of the cDNA sequences and predicts that the core block of repeats in the sp140 gene may be as small as the sum of these two *Hae*III fragments.

The genomic location of the sp140 gene was determined by *in situ* hybridization of recombinant cDNA plasmids to salivary gland polytene chromosomes. When pCt140.2 was labeled with [α-^{35}S]dCTP by nick translation and hybridized to a squashed preparation of chromosomes, autoradiography demonstrated that all detectable hybridization took place at a single site in the genome on polytene chromosome I (Fig. 7*A*). A comparison to published cytogenetic maps (58) indicated that the hybridization occurred at locus I-17-B (Fig. 7*B*). This location was confirmed at higher resolution by hybridization with biotinylated cDNA probes (data not shown). It is well known from cytological studies that this chromosome region forms an exceptionally large puff which attains maximal size in the middle of the prepupal stages of larval development (66).

The simplest conclusion drawn from our data was that chromosome locus I-17-B contained one copy of the sp140 gene. Most of the gene consists of contiguous blocks of 42-bp repeats similar to those found in pCt140.1 and pCt140.2.

Comparisons between sp140 and Other Secretory Proteins—Prior to this report, the partial structures of five secretory proteins have been deduced from the nucleotide sequences of

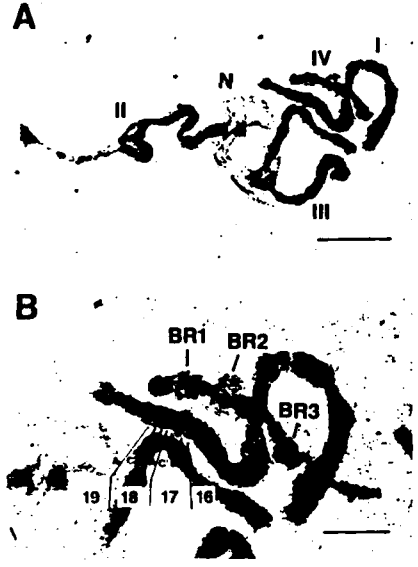

FIG. 7. **In situ hybridization of ^{35}S-labeled pCt140.1 to polytene chromosomes from the salivary gland of *C. tentans*.** *A*, inspection of all polytene chromosomes (*I, II, III, IV*) shows that silver grains are concentrated over a single locus in chromosome I. The nucleoli (*N*) are on chromosomes II and III. *B*, an enlargement of *A* showing BRs (*BR1, BR2*, and *BR3*) on chromosome IV and identification of the banding pattern on chromosome I (58). The silver grains are centered over chromosome band I-17-B. *Bars*: *A*, 250 μm; *B*, 50 μm.

TABLE I

Structural organization, amino acid composition, and common peptide sequences in secretory polypeptides from salivary glands of C. tentans

The amino acid composition and consensus oligopeptide sequences were derived from nucleotide sequence data obtained from published cDNA and genomic DNA clones. The ⊕ Pro ⊖ tripeptide motifs are underlined. Residues in alignment with similar or identical residues in the sp140 peptide are shown in boldface.

Secretory polypeptide	Major structural element	Most abundant residues											Shared oligopeptide sequences		Refs
		Lys	Arg	His	Glu	Asp	Pro	Ser	Cys	Gly	Ala	Asn	Location	Sequence	
		mol %													
spIa	82 residue tandem repeats (C + SR)ₙ	20	12	0	9	0	16	16	5	9	7	3	SR region	Lys-Pro-Ser-Lys-Gly-Ser-Lys Pro-Arg-Pro-Glu	60, 61
spIb	60–90 residue tandem repeats (C + SR)ₙ	14	20	0	2	0	10	20	6	9	6	0	SR1	Arg-Pro-Ser-Trp-Ser-Gly-Ile-Arg-Pro-Glu	62, 63
													SR2	Arg-Arg-Ser-Arg-Ser-Gly-Pro-Arg-Pro-Lys	
													SR3	Gly-Pro-Ser-Arg-Ser-Gly-Ser-Arg-Pro-Glu	
													SR4 & SR5	Lys-Arg-Ser-Lys-Ser-Gly-Ser-Arg-Pro-Glu	
spIc	79-residue tandem repeats (C + SR)ₙ	6	22	0	27	14	15	0	5	1	1	5	SR region	Arg-Pro-Glu-Arg-Pro-Glu-Arg-Pro-Glu	64
spId	74-residue tandem repeats (C + SR)ₙ	26	5	8	3	0	11	22	5	4	5	3	Junction of SR & C region	Lys-His-Ser-Lys-Pro-Ser	65
sp195	25 residue simple tandem repeats	20	0	0	12	8	12	0	8	0	4	12	Within repeat	Lys-Trp-Glu-Lys-Pro-Lys-Pro-Asp	27[1]
sp140	14 residue simple tandem repeats	28	0	0	4	0	13	13	0	30	3	8	Within repeat	Gly-Lys-Ser-Gly-Ser-Lys-Pro-Glu	Fig. 3

cDNA or genomic clones. Comparison of these structures reveals that sp140 and the other secretory proteins share several characteristics, although subtle differences exist (Table I). All secretory proteins contain repeated sequences of various lengths and complexity. It is interesting to note that sp140 is the smallest of these proteins and that it contains the shortest and most homogeneous repeats. The differences in repeat structure of these proteins result in differences in amino acid composition. Repeats in sp140 resemble repeats in the other secretory proteins in their abundance of charged residues, plus proline and serine. On the other hand, sp140 is unique in that glycine is the predominant amino acid and cysteine is virtually absent (except for the apparent substitution of nucleotides 131 and 133 within the most divergent repeat of pCt140.1). sp140 contains copies of the ⊕ Pro ⊖ tripeptide motif which has been observed in other secretory proteins. Whereas this motif occurs only once per repeat in sp140 and sp195, it occurs multiple times in the SR regions of the spIs. In essence, the simple repeat units of sp140 resemble tandem copies of spI SRs. In fact, SRs in spIb appear most similar to C-terminal half of sp140 repeats by virtue of a related heptameric sequence: Lys/Arg-Ser-Gly-Ser-Lys/Arg-Pro-Glu/Gly.

Developmental Expression of the sp140 Gene—In contrast to the spI family of secretory proteins, the maximum concentration of sp195 occurs in salivary glands of prepupae (stages 8–10 of the fourth larval instar; Ref. 29). We decided to determine whether or not the expression of sp140 was also developmentally regulated. Secretory proteins were extracted from the lumen of salivary glands of morphologically staged (30), fourth instar larvae and fractionated by gel electrophoresis. Protein blots made from such gels were stained (Fig. 8*A*), destained and subsequently incubated with the affinity-purified antipeptide antibody specific for sp140 (Fig. 4). A

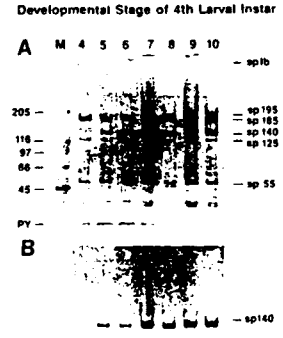

FIG. 8. **Developmental changes in the glandular content of sp140.** Salivary glands were removed from pools of fourth instar larvae from stages 4–10. Secretory polypeptides were extracted, electrophoresed, and blotted to nitrocellulose (Fig. 4). The blot was stained with Ponceau S (*A*), destained, reacted with the rabbit anti-(⊕Z peptide) antibody specific for sp140 and goat anti-rabbit antibody conjugated with alkaline phosphatase (*B*). Each lane contains two glands' worth of protein. Numbers of the left indicate the size (in kDa) of molecular mass markers (*M*) in a parallel lane of the gel. Numbers to the right identify secretory proteins (sp) by their apparent molecular mass (in kDa). *PY* indicates the position of Peronin Y, the tracking dye.

[1] V. L. Hill and S. T. Case, unpublished cDNA sequences.

reproducible developmental pattern of expression was observed (Fig. 8*B*): sp140 was usually undetectable at stage 4; it reached a detectable and nearly constant level during stages 5 and 6, beginning at stage 7, the level of sp140 again increased 2–3-fold exhibiting a maximum during stages 8–10.

To determine if the developmental pattern of sp140 expression was correlated with alterations in the steady-state concentration of its mRNA, Northern and dot-blots were made with RNA extracted from the salivary glands of staged larvae. Blots were hybridized with ^{32}P-labeled oligonucleotide C5.1 which is a specific probe for sp140 mRNA (Fig. 1*B*). The intensity of autoradiographic images made from these blots was compared by two-dimensional densitometry (Fig. 9). Over a series of developmental studies it was determined that there was a stage-specific pattern in the steady-state level of sp140 mRNA which mimicked the pattern observed for its protein: sp140 mRNA was initially detectable at stage 4, there was a 4-fold increase during stages 5 and 6; another increase occurred at stage 7 so that larvae at stages 8–10 typically had twice as much sp140 mRNA as stage 6 larvae. Quantitative dot-blots were used to compare the amount of probe which hybridized a gland's worth of RNA to known amounts of cDNA. We calculated that larvae in stages 8–10 (prepupae) contain as many as $2–7 \times 10^6$ sp140 mRNA molecules per secretory cell. This calculation was based upon the fact that each salivary gland contains 38 secretory cells (59) and the assumption that each sp140 mRNA molecule contains 85 copies (3600/42 nucleotides) of the 42-base repeat, each of which are similar enough to hybridize the oligonucleotide probe. This calculated abundance of sp140 mRNA is comparable to levels reported for the mRNAs of other secretory proteins (29).

Run-on transcription assays were done to compare the transcription rate of the sp140 gene during the fourth instar (Fig. 10). Nuclei were isolated from the salivary glands of staged larvae and allowed to continue RNA synthesis *in vitro* in the presence of [α-^{32}P]UTP. Nuclear RNA was extracted and hybridized to dot-blots containing the transcribed strand of sp140 cDNA. As larvae reached stage 5 there was about a 7-fold increase in the relative transcription rate of the sp140 gene, which coincided with the observed increase in the steady-state level of its mRNA (Fig. 9). This surge temporarily declined but increased again nearly 5-fold at stage 8 and essentially maintained a maximum rate between stages 8 and 10. We concluded that the expression of the sp140 gene is developmentally regulated during the fourth larval instar. The primary level of regulation most likely occurs at the level of transcription, although we cannot, at this time, rule out small contributions due to changes in mRNA stability.

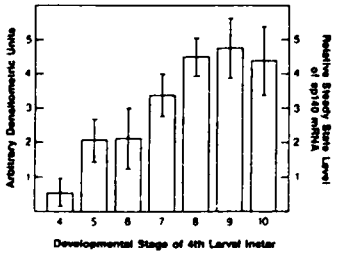

FIG. 9. **Developmental changes in the steady-state level of sp140 mRNA.** Total salivary gland RNA was extracted from pools of fourth instar larvae from stages 4–10. Equal quantities of staged RNA were used to make Northern blots (5–7 μg/lane) or dot-blots (0.5–1.5 μg/dot) that were hybridized with ^{32}P-labeled oligonucleotide C5.1 which is specific for the 3.6 kb sp140 mRNA (see Fig. 1*B*). Two-dimensional densitometry was used to compare the relative autoradiographic intensity of all bands and dots which reflect the relative level of sp140 mRNA. The *graph* represents the mean and standard deviation obtained over a series of four Northern blots plus three dot-blots. To quantify the results of dot blots, their images were also compared to those simultaneously obtained from a serial dilution of known quantities of pC140.1

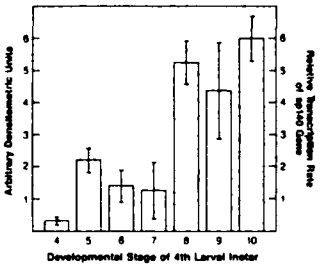

FIG. 10. **Developmental changes in the relative transcription rate of the sp140 gene.** Nuclei were isolated from the salivary glands of fourth instar larvae from stages 4–10. Run-on transcription assays were done and ^{32}P-labeled nuclear RNA was hybridized to dot-blots containing single-stranded cDNA inserts in M13mp18 DNA. Two-dimensional densitometry was used to compare the relative intensities of autoradiographic images. The *graph* displays the mean and standard deviation obtained from triplicate determinations.

Expression Classes of Secretory Protein Genes—We have identified a developmentally regulated gene for sp140 and mapped it to polytene chromosome band 1-17-B. The sp140 gene produces a 3.6-kb poly(A)$^+$ mRNA largely composed of tandem copies of 42-bp repeats. These repeats encode a peptide of 14 amino acids which shares a tripeptide motif common to all other known secretory proteins. This tripeptide is also part of a heptameric sequence (Table I) which is quite similar to SR regions in spIb. The conservation of this sequence may be important for our understanding of the evolution of this multigene family and may further define a site for intermolecular protein interactions during polymerization of secretory proteins into silk-like fibers (18, 19).

We propose that there are at least two expression classes of secretory protein genes. The spI genes represent a class of conditionally expressed larval genes. These genes as a class are generally expressed throughout all four larval instars (29, 56); however, concurrent expression of all spI genes apparently is not obligatory (57). Instead, individual genes may be able to respond to a variety of environmental effectors to ensure continuous synthesis of at least one of the structurally similar gene products. In contrast the sp140 and sp195 genes represent a separate class of developmentally regulated prepupal genes. Both are regulated at the level of transcription to go from undetectable to maximal levels of expression all within the fourth instar. Since both genes are maximally expressed during the prepupal stages of larval development we suggest that, similar to our hypothesis concerning sp195 (29), synthesis of sp140 may contribute to alterations in the microscopic and physical properties of the silk-like fibers spun by larvae constructing pupation tubes.

While the chromosomal distribution of the spI class of conditionally expressed larval genes in BRs may reflect their differential response to galactose, there doesn't seem to be a pattern to explain the distribution of developmentally regulated prepupal genes (sp140 in band 17 B on chromosome I, sp195 in BR1 on chromosome IV). Perhaps a pattern will emerge once additional secretory protein genes are identified and mapped. In the meantime, our results demonstrate that secretory protein genes exist outside of BRs and that their distribution extends to at least three of four chromosomes.

Acknowledgments—We wish to thank Dr. Stephen L. Eck for assistance with peptide syntheses, Dr. J. David Dignam for advice and assistance with immunization and blotting procedures, Dr. Thomas D. Dreesen for helping with the run-on nuclear transcription assays, Lizabeth L. Brumley for blots containing salivary gland poly(A)$^+$ RNA, and Dr. H. William Detrich for suggesting the use of Pyronin Y and Ponceau S. We are particularly grateful to our colleagues, Drs. Donald B. Sittman, and Susan E. Wellman for providing critical and helpful evaluations of this manuscript.

REFERENCES

1. Tolan, D. R., Niclas, J., Bruce, B. D., and Lebo, R. V. (1987) *Am. J. Hum. Genet.* **41**, 907–924
2. Gatti, R. A., Shaked, R., Mohandas, T. K., and Salser, W. (1987) *Am. J. Hum. Genet.* **41**, 654–667
3. Kent, R. B., Fallows, D. A., Geissler, E., Glaser, T., Emanuel, J. R., Lalley, P. A., Levenson, R., and Housman, D. E. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 5369–5373
4. Hawkins, J. W., Van Keuren, M. L., Piatigorsky, J., Law, M. L., Patterson, D., and Kao, F. T. (1987) *Hum. Genet.* **76**, 375–380
5. Emanuel, B. S., Sellinger, B. T., Gudas, L. J., and Myers, J. C. (1986) *Am. J. Hum. Genet.* **38**, 38–44
6. Neer, E. J., Michel, T., Eddy, R., Shows, T., and Seidman, J. G. (1987) *Hum. Genet.* **77**, 259–262
7. Ashley, P. L., Ellison, J., Sullivan, K. A., Bourne, H. R., and Cox, D. R. (1987) *J. Biol. Chem.* **262**, 15299–15301
8. Opdenakker, G., Cabeza-Arvelaiz, Y., Fiten, P., Dijkmans, R., Van Damme, J., Volckaert, G., Billiau, A., Van Elsen, A., Van der Schueren, B., Van den Berghe, H., and Cassiman, J.-J. (1987) *Hum. Genet.* **75**, 339–344
9. Tripputi, P., Emanuel, B. S., Croce, C. M., Green, L. G., Stein, G. S., and Stein, J. L. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3185–3188
10. Brown, D. T., Wellman, S. E., and Sittman, D. B. (1985) *Mol. Cell. Biol.* **5**, 2879–2886
11. Brown, D. T., Yang, Y.-S., and Sittman, D. B. (1988) *Mol. Cell. Biol.* **8**, 4406–4415
12. Meehan, R. R., Speed, R. M., Gosden, J. R., Rout, D., Hutton, J. J., Taylor, B. A., Hilkens, J., Kroezen, V., Hilgers, J., Adesnik, M., Friedberg, T., Hastie, N. D., and Wolf, C. R. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2662–2666
13. Maniatis, T., Fritsch, E. F., Lauer, J., and Lawn, R. M. (1980) *Annu. Rev. Genet.* **14**, 145–178
14. Ward, S., Burke, D. J., Sulston, J. E., Coulson, A. R., Albertson, D. G., Ammons, D., Klass, M., and Hogan, E. (1988) *J. Mol. Biol.* **199**, 1–13
15. Ingolia, T. D., Craig, E. A., and McCarthy, B. J. (1980) *Cell* **21**, 669–679
16. Garfinkel, M. D., Pruitt, R. E., and Meyerowitz, E. M. (1983) *J. Mol. Biol.* **168**, 765–789
17. Grond, C., Saiga, H., and Edstrom, J. E. (1987) in *Results and Problems in Cell Differentiation* (Henning, W., ed) Vol. 14, pp 69–80, Springer-Verlag, Berlin
18. Hamodrakas, S. J., and Kafatos, F. C. (1984) *J. Mol. Evol.* **20**, 296–303
19. Wieslander, L. W., Hoog, C., Hoog, J. O., Jornvall, H., Lendahl, U., and Daneholt, B. (1984) *J. Mol. Evol.* **20**, 304–312
20. Wallace, J. B., and Merritt, R. W. (1980) *Annu. Rev. Entomol.* **25**, 103–132
21. Walshe, B. M. (1950) *Proc. Zool. Soc. Lond.* **121**, 63–79
22. Edstrom, J. E., Rydlander, L., and Francke, C. (1980) *Chromosoma (Berl.)* **81**, 115–124
23. Edstrom, J. E., Sierakowska, H., and Burvall, K. (1982) *Dev. Biol.* **91**, 131–137
24. Galler, R., Saiga, H., Widmer, R. M., Lezzi, M., and Edstrom, J.-E. (1985) *EMBO J.* **11**, 2977–2982
25. Kao, W.-Y., and Case, S. T. (1985) *J. Cell Biol.* **101**, 1044–1051
26. Case, S. T. (1986) *Chromosoma (Berl.)* **94**, 483–491
27. Dreesen, T. D., Bower, J. R., and Case, S. T. (1985) *J. Biol. Chem.* **260**, 11824–11830
28. Dreesen, T. D., and Case, S. T. (1987) *Gene (Amst.)* **55**, 55–65
29. Dreesen, T. D., Lezzi, M., and Case, S. T. (1988) *J. Cell Biol.* **106**, 21–27
30. Ineichen, H., Meyer, B., and Lezzi, M. (1983) *Dev. Biol.* **98**, 278–286
31. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
32. Case, S. T., and Daneholt, B. (1978) *J. Mol. Biol.* **124**, 223–241
33. Kavenoff, R., and Zimm, B. H. (1973) *Chromosoma (Berl.)* **41**, 1–27
34. Wyss, C. (1982) *Exp. Cell Res.* **139**, 309–319
35. Thomas, P. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 5201–5205
36. Bailey, J. M., and Davidson, N. (1976) *Anal. Biochem.* **70**, 75–85
37. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517
38. Philippsen, P., Kramer, R. A., and Davis, R. W. (1978) *J. Mol. Biol.* **123**, 371–386
39. Sutcliffe, J. G. (1978) *Nucleic Acids Res.* **5**, 2721–2728
40. Kafatos, F. C., Jones, C. W., and Efstratiadis, A. (1979) *Nucleic Acids Res.* **7**, 1541–1552
41. Derksen, J. (1978) *Methods Cell Biol.* **17**, 133–139
42. Gall, J. G., and Pardue, M. L. (1971) *Methods Enzymol.* **21**, 470–480
43. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene (Amst.)* **33**, 103–119
44. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
45. Henikoff, S. (1984) *Gene (Amst.)* **28**, 351–359
46. Biggin, M. D., Gibson, T. J., and Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 3963–3965
47. Postell, J., and Kafatos, F. C. (1984) *Nucleic Acids Res.* **14**, 479–488
48. Merrifield, R. B. (1963) *J. Am. Chem. Soc.* **85**, 2149–2154
49. Kent, S. B. H. (1988) *Annu. Rev. Biochem.* **57**, 957–989
50. Rydlander, L. (1984) *Biochem. J.* **220**, 423–431
51. Burnette, W. N. (1981) *Anal. Biochem.* **112**, 195–203
52. Sabinovich, D., and Montelaro, R. C. (1986) *Anal. Biochem.* **156**, 341–347
53. Leary, J. J., Brigati, D. J., and Ward, D. C. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 4045–4049
54. Sass, H. (1980) *Chromosoma (Berl.)* **78**, 23–78
55. Wieslander, L. (1979) *J. Mol. Biol.* **134**, 347–367
56. Lendahl, U., and Wieslander, L. (1987) *Dev. Biol.* **121**, 130–138
57. Kao, W.-Y., and Case, S. T. (1986) *Chromosoma (Berl.)* **94**, 475–482
58. Beermann, W. (1952) *Chromosoma (Berl.)* **5**, 139–198
59. Case, S. T., and Daneholt, B. (1977) *Int. Rev. Biochem.* **15**, 45–77
60. Wieslander, L., Sumegi, J., and Daneholt, B. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 6956–6960
61. Case, S. T., and Byers, M. R. (1983) *J. Biol. Chem.* **258**, 7793–7799
62. Case, S. T., Summers, R. L., and Jones, A. G. (1983) *Cell* **33**, 555–562
63. Hoog, C., and Wieslander, L. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 5165–5169
64. Lendahl, U., and Wieslander, L. (1984) *Cell* **36**, 1027–1034
65. Sumegi, J., Wieslander, L., and Daneholt, B. (1982) *Cell* **30**, 579–587
66. Clever, U. (1962) *Chromosoma (Berl.)* **13**, 385–436

Gene 00449
Elsevier

Balbiani ring 3 in *Chironomus tentans* encodes a 185-kDa secretory protein which is synthesized throughout the fourth larval instar

(Recombinant DNA; in situ hybridization; partial cDNA sequence; Northern blots; antipeptide antibodies; immunoblots; salivary glands; polytene chromosomes)

Susan S. Dignam* and Steven T. Case

Department of Biochemistry, The University of Mississippi Medical Center, Jackson, MS 39216-4505 (U.S.A.)

Received by D.M. Skinner 18 September 1989
Revised 30 November 1989
Accepted 2 December 1989

see p. 2, 5

SUMMARY

We have continued to map and identify genes encoding a family of secretory proteins. These proteins are synthesized in larval salivary glands of the midge, *Chironomus tentans*, and assemble in vivo into insoluble silk-like threads. The genes for several secretory proteins exist in Balbiani rings (BRs) on salivary-gland polytene chromosomes. A randomly primed cDNA clone, designated pCt185, hybridized in situ to BR3 and was shown on Northern blots to originate from a salivary gland-specific 6-kb poly(A)$^+$ RNA. The partial cDNA sequence contained 483 nucleotides including one open reading frame (ORF) encoding 160 amino acids (aa). A striking feature of the ORF was the periodic distribution of cysteine residues (Cys-X-Cys-X-Cys-X_4-Cys) which occurred approximately every 22 aa. A cDNA-encoded 18-aa sequence was selected for chemical peptide synthesis. When affinity-purified antipeptide antibodies were incubated with a Western blot containing salivary-gland proteins they reacted specifically with a 185-kDa secretory protein (sp185). Developmental studies showed that sp185 and its mRNA were present in salivary glands throughout the fourth larval instar. Thus sp185 and a family of 1000-kDa secretory proteins are encoded by a class of genes that are expressed throughout the fourth instar. This contrasts with the developmentally regulated expression of the sp140 and sp195 genes whose expression is maximal during the prepupal stages of larval development.

INTRODUCTION

Cysteine (Cys) residues play an important role in determining protein conformation and function. Clusters of Cys residues coordinate metal binding (Williams et al., 1985; Evans and Hollenberg, 1988; Naqui et al., 1988). Cys pairs can provide intramolecular disulfide bonds which promote accurate protein folding required for conformational stability and enzymic activity (States et al., 1984; Eyerle and Schartau, 1985; Pace et al., 1988; Wetzel et al., 1988; Wilcox et al., 1988). Intermolecular disulfide bonds provide interactive sites for assembly of protein subunits into complexes (Fessler et al., 1985; Davis et al., 1988).

A group of large, Cys-containing secretory proteins are synthesized in larval salivary glands of the midge, *Chironomus*. Secretory proteins assemble in vivo into insoluble silk-like threads which larvae spin to construct underwater tubes for filter feeding and pupation (Walshe, 1950; Wallace and Merritt, 1980). The Cys-containing secretory proteins are related structurally in that they contain multiple tandem copies of aa sequences which include invariant Cys residues. For example, the spI family consists of four similar proteins (spIa, spIb, spIc and spId) with a molecular mass of about 1000 kDa (reviewed in Hamodrakas and Kafatos, 1984; Wieslander et al., 1984; Grond et al., 1987). Each spI is largely composed of tandem copies of a complex core repeat of 60–90 aa. These core repeats consist of two distinct regions. The C region contains 35–45 aa including four Cys, one Met and one Phe which are invariant. The SR consists of four to six direct repeats of a 6–12-aa sequence dominated by a tripeptide motif: Pro preceded by a positively charged aa (Lys or Arg) and followed by a negatively charged aa (Asp, Glu or phosphoSer). The designation used for this motif is ⊖Pro⊕. In addition to core repeats, the C end of each spI molecule contains several dispersed copies of an 18-aa sequence. These are known as Cys-1 repeats since each one contains a single invariant Cys residue (Grond et al., 1987). While the function of Cys residues in secretory proteins has yet to be determined, their distribution and evolutionary conservation suggest that intermolecular disulfide bonds may contribute to the assembly and insolubility of the secretory proteins in larval silk (Hamodrakas and Kafatos, 1984; Wieslander et al., 1984).

Salivary glands also contain a 195-kDa secretory protein (sp195; previously referred to as sp180 in Dreesen and Case, 1987; Dreesen et al., 1988). This protein is similar to spIs in that it contains tandemly repeated aa sequences (Dreesen et al., 1985). These are simple repeats of only 25 aa (Dreesen et al., 1985; Hill and S.T. Case, unpublished data), yet, each repeat contains two invariant Cys residues and one copy of the ⊕Pro⊖ tripeptide motif. Besides structural differences, sp195 and spIs differ markedly in their synthesis. Whereas spIs are synthesized throughout the larval instars, the synthesis of sp195 is limited to prepupal stages of larval development (Lendahl and Wieslander, 1987; Dreesen et al., 1988). The addition of sp195 to spIs may contribute to microscopic alterations observed in the structure of prepupal silk (Dreesen et al., 1988).

The Cys-containing secretory proteins share one other common feature: all their genes are located in cytological structures known as BRs. BRs are enormous tissue-specific puffs located on polytene chromosomes in salivary glands (Beermann, 1972). BR1 contains the genes encoding spIa and sp195, BR2 contains the genes encoding spIb and spIc, and BR6 contains the gene encoding spIc (Dreesen et al., 1985; Case, 1986, and references cited therein).

The aim of the present study was to identify a gene in BR3 of *C. tentans*. With the aid of cDNA and antipeptide antibody probes, we wanted to learn whether or not BR3 also contained a gene encoding a secretory protein and determine when is this gene was expressed during the fourth instar.

RESULTS AND DISCUSSION

(a) BR3 contains a gene for a 6-kb poly(A)$^+$ RNA

Salivary-gland secretory proteins are tissue-specific and most of their genes reside in BRs. To identify additional genes encoding secretory proteins we continued to screen cDNA clones that were randomly primed from salivary gland RNA (Dreesen et al., 1985) by a combination of in situ hybridization to polytene chromosomes and Northern-blot analyses of salivary-gland RNA. One cDNA clone, designated here as pCt185, hybridized exclusively to BR3 on polytene chromosome IV (Fig. 1). Hybridization of pCt185 to a Northern blot of salivary-gland RNA demonstrated that this cDNA originated from a 6-kb poly(A)$^+$ RNA (Fig. 2A). We were unable to obtain hybridization of pCt185 to Northern blots containing hybridizable bands of RNA from other larval tissues (data not shown). Therefore, this cDNA might have been derived from a tissue-specific mRNA from BR3. This was of particular interest since BR3 remained the only BR in *C. tentans* without an identified gene product.

(b) The *BR3* gene encodes a 185-kDa secretory protein

A *Hae*III fragment of pCt185 was subcloned into bacteriophage M13mp8 DNA to create strand-specific hybridization probes to orient the transcriptional polarity of the cDNA insert (Figs. 2B and 3). M13 templates and deletion constructs (Fig. 3) were used to obtain the complete 483-nt sequence of this partial cDNA (Fig. 4). The mRNA sequence contained only one ORF encoding a sequence of 160 aa (Fig. 4); the other two reading frames contained stop codons.

An oligopeptide (corresponding to codons between nt 228–281 in Fig. 4) was synthesized and used to raise rabbit antipeptide antibodies for the immunological identification of a *BR3* gene product. When affinity-purified antibodies were incubated with a Western blot containing an extract of total salivary-gland proteins they reacted selectively with a 185-kDa protein (Fig. 5). The specificity of this reaction was confirmed by the loss of immunoreactivity if

Correspondence to: Dr. S.T. Case, Department of Biochemistry, The University of Mississippi Medical Center, 2500 North State Street, Jackson, MS 39216-4505 (U.S.A.) Tel. (601)984-1518; Fax (601)984-1262.
* Present address: Department of Biochemistry, Medical College of Ohio, C.S. 10008, Toledo, OH 43699 (U.S.A.) Tel. (419)381-4131.

Abbreviations: aa, amino acid(s); bp, base pair(s); BR, Balbiani ring; BSA, bovine serum albumin; C region, constant region of an spI core repeat; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; PA, polyacrylamide; SET, 0.15 M NaCl/2 mM EDTA/30 mM Tris·HCl pH 8.0; spI, a 1000-kDa secretory protein; sp'x', other secretory proteins with 'x' indicating the apparent size in kDa; SR region, subrepeat region of an spI core repeat; X, any aa; ⊖, a positively charged aa (Arg or Lys); ⊖, a negatively charged aa (Asp, Glu, or phosphoSer).

0378-1119/90/$03.50 © 1990 Elsevier Science Publishers B.V. (Biomedical Division)

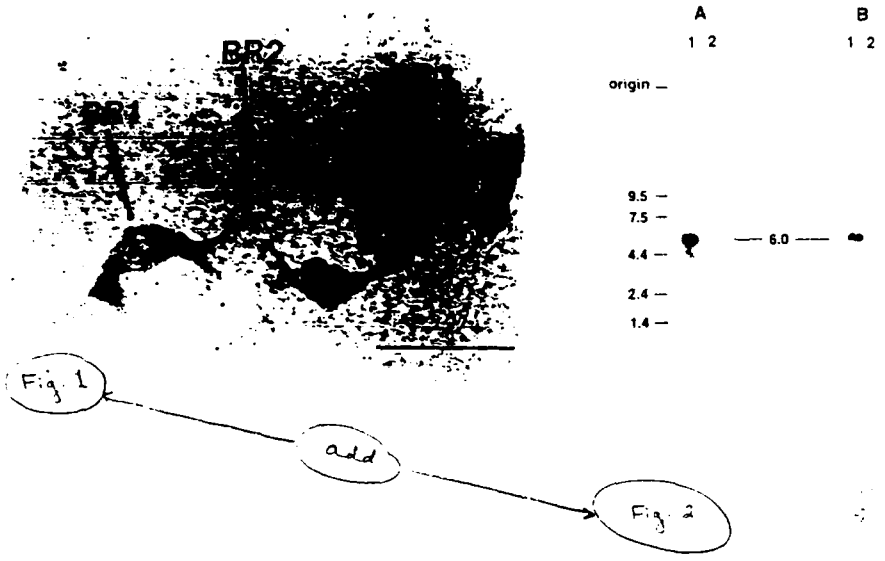

Fig. 1. Hybridization of ^{35}S-labeled cDNA plasmid pCt185 to salivary gland polytene chromosomes in situ. Plasmid DNA was labeled with [α-^{35}S]dCTP by nick-translation and hybridized in situ to squashed preparations (Derksen, 1978) of salivary gland polytene chromosomes at 65°C for 24 h in 4 × SET, containing 0.1% K pyrophosphate and 500 μg of heparin/ml. The posthybridization rinse included 1 h in 0.1 × SET at 65°C. Autoradiographic detection of hybridization (3-week exposure) and staining of polytene chromosomes were done as described (Gall and Pardue, 1971). The photomicrograph shows chromosome IV with BRs (BR1, BR2, and BR3) labeled. Silver grains are seen over BR3. Bar = 50 μm.

Fig. 2. Hybridization of cDNA and oligo probes to Northern blots of salivary gland RNA. Procedures described without specific literature citations were similar to protocols described in Maniatis et al. (1982). Salivary-gland RNA was extracted (Case and Daneholt, 1978) and, where indicated, chromatographed over oligo(dT) cellulose to obtain poly(A)$^+$ RNA. Northern blots (Thomas, 1980) of salivary-gland RNA were made from denaturing 0.75% agarose gels containing methylmercury hydroxide (Bailey and Davidson, 1976). *RNA markers were obtained from Bethesda Research Laboratories* (Catalogue No. 5620A). Hybridizations were done in 4 × SET containing 0.1% SDS/0.1M K-pyrophosphate/500 μg of heparin/ml. Plasmid probes were labeled with [α-^{32}P]dCTP and hybridized for 24–48 h in 4 × SET at 65°C, rinsed twice for 30 min in 2 × SET at room temperature, and 1 h in 1 × SET at 65°C. Each oligo probe (*see legend to Fig. 3 for their sequences and* orientation within the cDNA) was end-labeled with [γ-^{32}P]ATP and hybridized in 4 × SET at a temperature determined experimentally to be T_m − 5°C (this value ranged from 53°C for C2 to 65°C for C1). Post-hybridization washes for oligo probes included four rinses in 2 × SET at room temperature and a 5-min rinse in 2 × SET at hybridization temperature. Autoradiograms show: lane A) 1 μg each of (1) poly(A)$^+$ and (2) poly(A)$^-$ salivary-gland RNA hybridized with ^{32}P-labeled oligos (1) C2 and (2) CC2. Numbers on the left margin indicate the position and size (in kb) of RNA markers that were run in parallel lanes of the gel and located by staining with ethidium bromide.

the antipeptide antibodies were preincubated with the synthetic peptide. To determine if the 185-kDa protein was a secretory protein, these same antibodies were incubated with a blot containing an extract of proteins from the lumen of salivary glands. A protein with similar electrophoretic mobility was detected (Fig. 5) demonstrating that the 185-kDa protein was secreted into the glandular lumen. We concluded that the ORF in the pCt185 cDNA sequence represents a 160-aa portion (approx. 10%) of sp185.

Several features of the aa sequence encoded in pCt185 *were reminiscent of other secretory proteins.* The most striking feature of this sequence was its content and periodic distribution of Cys residues: the pattern Cys-X-Cys-X-Cys-X_6-Cys occurred almost every 22 aa. Thus far, this pattern seems unique to sp185 and it is not part of a repeating aa sequence. All other BR-encoded secretory proteins contain tandemly repeated sequences which include conserved Cys residues. We also noted that there were three

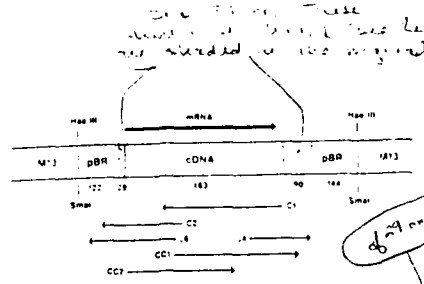

Fig. 3. A diagram of the transcriptional polarity and DNA sequencing strategy used for the cDNA insert of pCt185. The construction of randomly primed cDNA clones in pBR322 was as described (Dreesen et al., 1985). cDNA was inserted *into the PstI site of pBR322* using dT·dA homopolymeric tails. A *Hae*III fragment of pCt185, including all 483 bp of cDNA, flanking homopolymeric tails (shaded regions) and adjacent *sequences from* pBR322, was blunt-end ligated into the *Sma*I site of phage M13mp8 (Yanisch-Perron et al., 1985). The arrow above the insert indicates the direction of transcription which was derived by hybridization of strand-specific M13 (data not shown) and oligo (Fig. 2B) probes to Northern blots of salivary-gland RNA. Arrows below the insert show the direction and extent of dideoxyribonucleotide-terminated sequencing reactions (Sanger et al., 1977). Reactions were done on full-length inserts using oligo primers (C1, 5′-CCATGTCTTGGCTCCACATCC-3′; C2, 5′-GGACTAGTACTTGGATTATTTTTGC-3′; CC1, 5′-CGATAAA-CCATCATGCGAATGC-3′; and CC2, 5′-GGAAAAATGCAAGTCACCAAGAC-3′) or on *deletion constructs* (J4, J6) made with exonuclease III (Henikoff, 1984) and sequenced using an M13 universal primer.

copies of a putative ⊖Pro⊖ tripeptide motif (Fig. 4). Finally, the most abundant aa (12%, Lys, 9%, Pro, 9%, Gly, 6%, Asn, 6%, Gln, 5%, Ser, and 5% Glu) are found in core repeats of sp185 (Dignam et al., 1989).

(c) Developmental expression of the gene encoding sp185

We recently learned that the expression of some, but not all, genes encoding secretory proteins is regulated developmentally during the fourth larval instar (Dreesen et al., 1988; Dignam et al., 1989). Thus, we decided to examine the levels of sp185 during development. When glandular protein was examined by immunoblotting, we found that sp185 was detectable at stages 3 through 10 of the fourth instar (Fig. 6). The apparent gradual increase in the glandular content of sp185 is less than the increase in size of the salivary gland during these stages of growth.

Northern blots were prepared to compare the steady-state level of sp185 mRNA to the developmentally regulated patterns reported for sp140 (Dignam et al., 1989) and sp195 (Dreesen et al., 1988) mRNAs. Each lane in the blot contained an equal quantity of RNA extracted from salivary glands of staged larvae. The blot was probed sequentially with end-labeled oligos specific for sp140 and sp195 mRNAs and nick-translated pCt185 (Fig. 7). At all stages

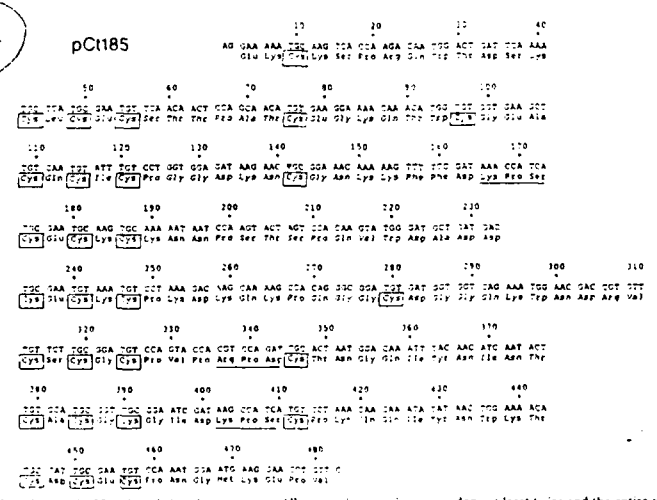

Fig. 4. The nt sequence of cDNA in pCt185 and its deduced aa sequence. All sequencing reactions were done at least twice and the entire sequence of both strands was determined independently. The nt sequences were compiled and analyzed using the programs of Pustell and Kafatos (1986) which were purchased from International Biotechnologies, Inc. Numbers above each line refer to the nt marked by an asterisk. The aa sequence of the only ORF is displayed; the other two reading frames contained stop codons. All Cys residues are boxed. The presentation of the sequence aligns the Cys-X-Cys-X-Cys pentapeptide which occurs nearly every 22 aa. Underlined tripeptides may be similar to the ⊖Pro⊖ tripeptide motif [(Lys,Arg)-Pro-(Glu,Asp,phosphoSer)] found in other secretory proteins (for summary, see Dignam et al., 1989). *The sequence reported in this paper* has been submitted to the Genbank EMBL Data Bank with accession number M24160.

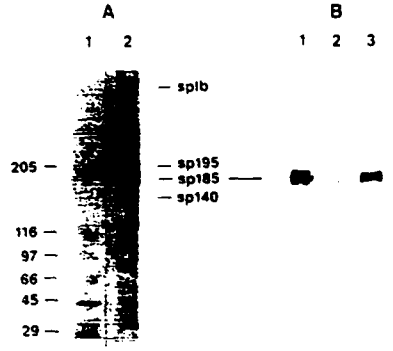

Fig. 5. Immunological identification of a protein encoded by pCt185 cDNA. Secretory proteins were extracted from salivary glands in 6 M guanidine·HCl, reduced, alkylated and fractionated by gel electrophoresis on 3–20% concave exponential gradients of PA containing SDS, as described (Kao and Case, 1985). The sample application buffer contained pyronin Y as the tracking dye which was run to the bottom of the gel. All gels contained a mixture of marker proteins (myosin, 205 kDa; β-galactosidase, 116 kDa; phosphorylase B, 97 kDa; BSA, 66 kDa; ovalbumin, 45 kDa; and carbonic anhydrase, 29 kDa) in a parallel lane. Western blots were made by electrophoretic transfer of proteins to nitrocellulose (Burnette, 1981), and stained with Ponceau S (Sabinovich and Montelaro, 1986). Blots were photographed, destained and used for immunoblotting. A cDNA-encoded aa sequence (nt 228–281 in Fig. 4) was selected for oligopeptide synthesis on an Applied Biosystems Model 430A Peptide Synthesizer. The peptide sequence was: (NH_2)-Asp-Asp-Cys-Glu-Cys-Lys-Cys-Pro-Lys-Asp-Lys-Gln-Lys-Pro-Gln-Gly-Gly-Cys-$(COOH)$. The peptide was cleaved from its support resin with hydrofluoric acid, chromatographed in 10 mM triethylamine bicarbonate pH 7.5, over a column of Sephadex G-50, lyophilized, reduced for 1 h in 10 mM dithiothreitol, alkylated for 1 h in 100 mM iodoacetamide and chromatographed in H_2O over a column of Sephadex G-25. The peptide was coupled to BSA with glutaraldehyde and injected into rabbits to obtain polyclonal antipeptide antisera. Antipeptide antibodies were immunoaffinity-purified (Dreesen and Case, 1987) by sequentially chromatographing antisera over columns containing either BSA or reduced and alkylated peptide coupled covalently to Affi-Gel 10 (BioRad Laboratories). Primary rabbit antipeptide antibodies were detected with secondary goat anti-rabbit antibodies coupled to alkaline phosphatase (Leary et al., 1983). (Panel A) Strips of nitrocellulose stained with Ponceau S with (lane 1) M, standards and (lane 2) total salivary gland proteins that were fractionated on PA gels and electrophoretically blotted to the membrane. (Panel B) Destained nitrocellulose strips that were incubated with a 1:20 dilution of affinity purified rabbit antipeptide antibody without (lanes 1 and 3) or with (lane 2) 175 μM reduced and alkylated synthetic oligopeptide used as the immunogen. Whereas lanes 1 and 2 in panel B contain total salivary gland protein, lane 3 contains an extract of secretory proteins from the lumen of salivary glands. Numbers on the left margin indicate the size (in kDa) and location of protein markers. Identifiable secretory proteins (sp140, sp185, sp195 and sp1b) are labeled.

examined, sp185 mRNA was detectable and its relative level varied generally less than fourfold. There was one notable exception: in two out of seven experiments (for example, see

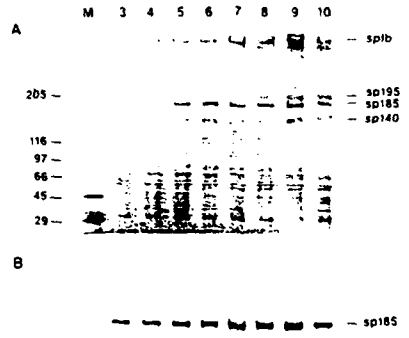

Fig. 6. Changes in the glandular content of sp185 during stages of the fourth larval instar. Larvae were raised as described (Dreesen et al., 1988) except that they were shifted to 20°C with a 16-h light/8-h dark cycle prior to reaching the fourth instar to prevent their entering diapause. Individual live larvae were staged based upon the morphology and orientation of imaginal discs (Ineichen et al., 1983). Salivary glands were dissected manually under a stereo microscope. Those used for developmental studies of mRNA and protein levels were stored in 70% ethanol at −20°C. Glands from larvae at similar stages were pooled. (Panel A) Protein markers (lane M) and two glands' worth of secretory proteins from staged larvae (stages 3 though 10) were fractionated by electrophoresis on PA gels, blotted to nitrocellulose and stained with Ponceau S. (Panel B) The blot in panel A was destained, reacted with rabbit antipeptide antibodies and alkaline phosphatase-conjugated goat anti-rabbit antibodies as described in Fig. 5. Numbers to the left indicate the size (kDa) and position of marker proteins. Numbers to the right indicate identifiable secretory proteins.

Fig. 7), stage-4 larvae exhibited more sp185 mRNA than any of the other stages. Why this occurred only occasionally is unclear; however, subsequent removal of the mRNA probes and rehybridization of the blots with cloned rRNA probes (data not shown) indicated that each lane did, in fact, contain equal amounts of RNA. Furthermore, the relatively constant pattern observed for sp185 mRNA contrasts markedly with the developmentally regulated patterns of sp140 and sp195 mRNAs; their steady-state concentrations changed dramatically within the fourth instar going from undetectable to maximal levels that were attained between stages 8–10 (the prepupal stages).

(d) sp185 mRNA is abundant in secretory cells

The steady-state concentration of one secretory protein's mRNA cannot be compared to another simply on the basis

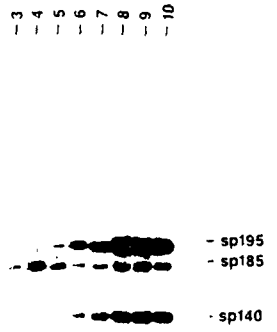

Fig. 7. Changes in the steady-state levels of secretory protein mRNAs during stages of the fourth larval instar. Salivary gland RNA was extracted from pools of staged glands. Each lane contains 7.2 μg of RNA which were fractionated on a denaturing agarose gel. RNAs were blotted onto a Nytran membrane and sequentially hybridized with ^{32}P-labeled oligo probes specific for sp140 (C5.1 in Dignam et al., 1989) and sp195 (TD1544A in Dreesen et al., 1985; 1988) mRNAs and plasmid probe pCt185 for sp185 mRNA. The blot was autoradiographed after each round of hybridization to monitor the specificity of each probe.

of a comparison of the autoradiographic intensity of bands in Fig. 7 for at least three reasons: (1) the hybridization probes had different specific activities; (2) mRNAs encoding sp140 (Dignam et al., 1989) and sp195 (Dreesen et al., 1985) contain tandemly repeated sequences, whereas sp185 may not (Fig. 4); and (3) the overall homogeneity of repeated sequences within sp140 and sp195 mRNAs is unknown. To calculate the cellular concentration of sp185 mRNA, quantitative dot blots were made (Kafatos et al., 1979; Case, 1986). By two-dimensional scanning of timed autoradiograms (Dignam et al., 1989), we measured the amount of hybridization obtained when oligo C2 was hybridized simultaneously to samples of salivary gland RNA and a serial dilution of a known quantity of pCt185 (data not shown). During stages 8–10, RNA from a single salivary gland hybridized as many copies of the C2 probe as did 1–3 ng of plasmid DNA. From this result we calculated that prepupae contain between $0.4–1.4 \times 10^7$ molecules of sp185 mRNA per secretory cell. This calculation was based upon the following assumptions: (1) the size of pCt185 is 4964 bp (Fig. 3); (2) pCt185 hybridized only one copy of the oligo probe (Fig. 4); (3) the oligo hybridized exclusively to sp185 mRNA (Fig. 2B); (4) only one copy of the probe hybridized each sp185 mRNA molecule; and (5) there are 8 secretory cells per salivary gland (Case and Daneholt, 1977). Thus, in spite of the apparent differences in autoradiographic signal, the steady-state concentration of sp185 mRNA is comparable to maximal levels measured for mRNAs encoding other secretory proteins (Dreesen et al., 1988; Dignam et al., 1989).

(e) All BRs contain secretory protein genes

A variety of cytological and biochemical data led earlier workers to propose that BRs contain the most actively transcribed genes in Chironomid salivary glands and that these genes encoded abundant mRNAs for secretory proteins (Case and Daneholt, 1977; Grossbach, 1977). The results of experiments described in this paper complete the identification of at least one secretory protein gene for each of the four BRs found on the polytene chromosomes in salivary glands of *C. tentans* (Table I). However, the location of secretory protein genes is not limited to BRs; the gene encoding sp140 is located in chromosome region 1-17-B (Dignam et al., 1989).

TABLE I

Summary of the identification, chromosomal location and expression of genes encoding secretory proteins in salivary glands of *Chironomus tentans*

Locus	Chromosome	Gene product	Apparent molecular size of protein (kDa)	Expression during fourth instar
BR1	IV	sp1a	approx. 1000	Throughout
		sp195	195	Prepupal stages
BR2	IV	sp1b	approx. 1000	Throughout
		sp1d	approx. 1000	Throughout
BR3	IV	sp185	185	Throughout
BR6	III	sp1c	approx. 1000	Inducible throughout
17-B	I	sp140	140	Prepupal stages

(f) Larval and prepupal secretory proteins

Secretory proteins can be grouped broadly into two expression classes: larval and prepupal (Table I). The spls represent the larval class because at least one of them is synthesized throughout the fourth instar (Lendahl and Wieslander, 1987; Dreesen et al., 1988). This pattern of expression and structural similarities of these proteins strengthened the notion that the structural backbone of larval silk could be either a homo- or heteropolymer of spls (Kao and Case, 1986). In fact, spls can assemble in vitro into fibrous secretory protein complexes (Wellman and Case, 1989). Results presented here would add sp185 to the class of larval proteins and suggest its involvement in the process of forming larval silk primarily for the construction of feeding tubes. In contrast, sp140 and sp195 constitute a class of developmentally regulated prepupal proteins (Dreesen et al., 1988; Dignam et al., 1989). Since the developmentally regulated synthesis of sp140 and sp195 coincides with microscopic changes observed in prepupal silk fibers in vivo, it has been proposed (Dreesen et al., 1988) that the addition and/or substitution of these proteins in larval silk leads to a structurally modified silk required for the construction of specialized pupation tubes.

(g) Conclusions

BR3 in the salivary glands of *C. tentans* contains a tissue-specific gene which is transcribed into a 6-kb mRNA for sp185. The gene encoding sp185 is expressed throughout the fourth larval instar and the steady-state level of its mRNA is comparable to levels of mRNAs for other secretory proteins in larval salivary glands. Like other secretory proteins, sp185 contains a unique pattern of Cys residues.

The abundance and nonrandom distribution of Cys in secretory proteins suggest that Cys residues play an important role in the assembly of secretory protein complexes and/or contribute towards their insolubility in vivo. The challenge which lies ahead is to learn the distribution of secretory proteins within the architecture of assembled complexes and identify the sites and nature of protein–protein interactions which take place between them.

ACKNOWLEDGEMENTS

We thank Lily Yang for help with the in situ hybridizations, Lizabeth Brumley for providing Northern blots with poly(A)$^+$ RNA and J. David Dignam for valuable suggestions during the purification of oligopeptides and immunoblotting experiments. We are particularly grateful to our colleagues, Susan E. Wellman and Donald B. Sittman, for reading and criticizing this manuscript. This research was supported by Office of Naval Research Contract No. N00014-87-K-0387 awarded to S.T. Case.

REFERENCES

Bailey, J.M. and Davidson, N.: Methylmercury as a reversible denaturing reagent for agarose gel electrophoresis. Anal. Biochem. 70 (1976) 75–85.

Beermann, W.: Chromosomes and genes. In Beermann, W. (Ed.), Results and Problems in Cell Differentiation, Vol. 4. Springer Verlag, Berlin, 1972, pp. 1–33.

Burnette, W.N.: 'Western blotting': electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radiolabeled protein(s). Anal. Biochem. 112 (1981) 195–203.

Case, S.T.: Correlated changes in steady-state levels of Balbiani ring mRNAs and secretory polypeptides in salivary glands of *Chironomus tentans*. Chromosoma 94 (1986) 483–491.

Case, S.T. and Daneholt, B.: Cellular and molecular aspects of genetic expression in *Chironomus* salivary glands. Int. Rev. Biochem. 15 (1977) 45–77.

Case, S.T. and Daneholt, B.: The size of the transcription unit in Balbiani ring 2 of *Chironomus tentans* derived from an analysis of the primary transcript and 75S RNA. J. Mol. Biol. 124 (1978) 223–241.

Davis, A.C., Roux, K.H. and Schulman, M.J.: On the structure of polymeric IgM. Eur. J. Immunol. 18 (1988) 1001–1008.

Derksen, J.: Cytological analysis of *Drosophila* polytene chromosomes. Methods Cell Biol. 17 (1978) 133–139.

Dignam, S.S., Yang, L., Lezzi, M. and Case, S.T.: Identification of a developmentally regulated gene for a 140-kDa secretory protein in salivary glands of *Chironomus tentans* larvae. J. Biol. Chem. 264 (1989) 9444–9452.

Dreesen, T.D. and Case, S.T.: A peptide-reactive antibody to a Balbiani ring gene product: immunological evidence that a 6.5-kb RNA in *Chironomus tentans* salivary glands is mRNA for a 180-kDa non-fibrous component of larval secretion. Gene 55 (1987) 55–65.

Dreesen, T.D., Bower, J.R. and Case, S.T.: A second gene in a Balbiani ring: *Chironomus* salivary glands contain a 6.5-kb poly(A)$^+$ RNA that is transcribed from a hierarchy of tandem repeated sequences in Balbiani ring 1. J. Biol. Chem. 260 (1985) 11824–11830.

Dreesen, T.D., Lezzi, M. and Case, S.T.: Developmentally regulated expression of a Balbiani ring 1 gene for a 180-kDa secretory polypeptide in *Chironomus tentans* salivary glands before larval/pupal ecdysis. J. Cell Biol. 106 (1988) 21–27.

Evans, R.M. and Hollenberg, S.M.: Zinc fingers: gilt by association. *Cell* 52 (1988) 1–3.

Eyerle, F. and Schartau, W.: *Hemocyanins in spiders, XX.* Sulfhydryl groups and disulfide bridges in subunit d of *Eurypelma californicum* hemocyanin. Hoppe Seyler's Z. Biol. Chem. 366 (1985) 403–409.

Fessler, J.H., Doege, K.J., Duncan, K.G. and Fessler, L.I.: Biosynthesis of collagen. J. Cell Biochem. 28 (1985) 31–37.

Gall, J.G. and Pardue, M.L.: Nucleic acid hybridization in cytological preparations. Methods Enzymol. 21 (1971) 470–480.

Grond, C., Saiga, H. and Edstrom, J.-E.: The sp-I genes in the Balbiani rings of *Chironomus* salivary glands. In Henning, W. (Ed.), Results and Problems in Cell Differentiation, Vol. 14. Springer-Verlag, Berlin, 1987, pp. 69–80.

Grossbach, U.: The salivary gland of *Chironomus* (Diptera): a model system for the study of cell differentiation. In Beermann, W. (Ed.), Results and Problems in Cell Differentiation, Vol. 8. Springer-Verlag, Berlin, 1977, pp. 147–196.

Hamodrakas, S.J. and Kafatos, F.C.: Structural implications of primary sequences from a family of Balbiani ring-encoded proteins in *Chironomus*. J. Mol. Evol. 20 (1984) 296–303.

Henikoff, S.: Unidirectional deletion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28 (1984) 351–359.

Ineichen, H., Meyer, B. and Lezzi, M.: Determination of the developmental stage of living fourth instar larvae of *Chironomus tentans*. Dev. Biol. 98 (1983) 278–286.

Kafatos, F.C., Jones, C.W. and Efstratiadis, A.: Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucleic Acids Res. 7 (1979) 1541–1552.

Kao, W.-Y. and Case, S.T.: A novel giant secretion polypeptide in *Chironomus* salivary glands: implications for another Balbiani ring gene. J. Cell Biol. 101 (1985) 1044–1051.

Kao, W.-Y. and Case, S.T.: Individual variations in the content of giant secretory polypeptides in salivary glands of *Chironomus* Chromotoma 94 (1986) 375–382.

Leary, J.J., Brigati, D.J. and Ward, D.C.: Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose: Bio-blots. Proc. Natl. Acad. Sci. USA 80 (1983) 4045–4049.

Lendahl, U. and Wieslander, L.: Balbiani ring (BR) genes exhibit different patterns of expression during development. Develop. Biol. 121 (1987) 130–138.

Maniatis, T., Fritsch, E.F. and Sambrook, J.: Molecular Cloning. A *Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring* Harbor, NY, 1982.

Naqui, A., Powers, L., Lundeen, M., Constantinescu, M. and Chance, B.: On the environment of zinc in beef heart cytochrome *c* oxidase: an x-ray absorption study. J. Biol. Chem. 263 (1988) 12342–12345.

Pace, C.N., Grimsley, G.R., Thompson, J.A. and Barnett, B.J.: Conformational stability and activity of ribonuclease T1 with zero, one and two intact disulfide bonds. J. Biol. Chem. 263 (1988) 11820–11825.

Pustell, J. and Kafatos, F.C.: A convenient and adaptable microcomputer environment for DNA and protein sequence manipulation and analysis. Nucleic Acids Res. 14 (1986) 479–488.

Sabinovich, D. and Montelaro, R.C.: Reversible staining and peptide mapping of proteins transferred to nitrocellulose after separation by sodium dodecylsulfate-polyacrylamide gel electrophoresis. Anal. Biochem. 156 (1986) 341–347.

Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

States, D.J., Dobson, C.M. and Karplus, M.: A new two-disulphide intermediate in the refolding of reduced bovine pancreatic trypsin inhibitor. J. Mol. Biol. 174 (1984) 411–418.

Thomas, P.: Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77 (1980) 5201–5205.

Wallace, J.B. and Merritt, R.W.: Filter-feeding ecology of aquatic insects. Annu. Rev. Entomol. 25 (1980) 103–132.

Walshe, B.M.: *The feeding habits of certain Chironomid larvae* (sub-family tendipedinae). Proc. Zool. Soc. Lond. 121 (1950) 63–79.

Wellman, S.E. and Case, S.T.: Disassembly and reassembly in vitro of complexes of secretory proteins from *Chironomus tentans* salivary glands. J. Biol. Chem. 264 (1989) 10878–10883.

Wetzel, R., Perry, L.J., Baase, W.A. and Becktel, W.J.: Disulfide bonds and thermal stability of T4 lysozyme. Proc. Natl. Acad. Sci. USA 85 (1988) 401–405.

Wieslander, L., Hodg, C., Hödg, J.-O., Jornvall, H., Lendahl, U. and Daneholt, B.: Conserved and non-conserved structures in the secretory proteins encoded in the Balbiani ring genes of *Chironomus tentans*. J. Mol. Evol. 20 (1984) 304–312.

Wilcox, W.C., Long, D., Sodora, D.L., Eisenberg, R.J. and Cohen, G.H.: The contribution of cysteine residues to antigenicity and extent of processing of herpes simplex virus type 1 glycoprotein D. J. Virol. 62 (1988) 1941–1947.

Williams, J., Moreton, K. and Goodearl, A.D.: Selective reduction of a disulphide bridge in hen ovotransferrin. Biochem. J. 228 (1985) 661–665.

Yanisch-Perron, C., Vieira, J. and Messing, J.: Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33 (1985) 103–119.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1989 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 264, No. 18, Issue of June 25, pp. 10878-10883, 1989
Printed in U.S.A.

Disassembly and Reassembly *in Vitro* of Complexes of Secretory Proteins from *Chironomus tentans* Salivary Glands*

(Received for publication, January 10, 1989)

Susan E. Wellman‡ and Steven T. Case§

From the Department of Biochemistry, University of Mississippi Medical Center, Jackson, Mississippi 39216-4505

The secretory proteins of *Chironomus tentans* larvae form insoluble fibers that are spun into threads used to construct underwater feeding and pupation tubes. We began *in vitro studies of the mechanism of assembly* into fibers, the structure of the assembled proteins, and the contribution of individual proteins to the assembled structure. From measurements of turbidity and electron micrographs, we observed that the secretory proteins were isolated as complexes. These complexes are most likely at initial stages of assembly; further assembly into insoluble fibers must occur *in vivo*. Denaturation and reduction disrupted the complexes, and removal of the denaturing and reducing agents resulted in reassembly of the complexes. The circular dichroic spectrum of the complexes indicated that the assembled proteins had the tertiary structure $\alpha + \beta$. The largest secretory proteins were purified and shown to have both similar morphology, using electron microscopy, and a similar dichroic spectrum to that of the native complexes. We concluded that the large secretory proteins form the fibrous backbone of the complexes that we observe.

Many arthropods produce and secrete filamentous proteins. Studies of these fibrous proteins, for example the fibroins of silkworms and spiders, have contributed significantly to our understanding of the relationship between amino acid sequence and protein conformation, as well as between conformation and mechanical properties such as elasticity or flexibility (Lucas *et al.*, 1960; Anderson, 1970). Larvae of the fly *Chironomus tentans* secrete proteins which are spun into insoluble fibrous material in their aquatic environment. The necessity of assembly and function under water imposes constraints on these proteins which may be reflected in their primary sequences and their tertiary and quaternary structure. Studies of these proteins should provide further information on the relationship between arrangements of amino acids and specific physicochemical and mechanical properties of proteins.

Salivary glands of the aquatic larvae of C. tentans contains a family of secretory proteins (SPs).[1] At least 12 SPs are found in the lumen of the salivary gland, which is the site of storage of the proteins to be secreted (Grossbach, 1977). SPs are secreted as silken threads that larvae spin under water to *construct feeding and pupation tubes (Wallace and Merritt,* 1980). Some of these proteins are spun into fibers, while others of the proteins may be used as a gluelike substance to hold the fibers together or to confer adhesiveness to the fibers (Dreesen *et al.*, 1988). Partial amino acid sequences have been inferred from cDNA or genomic DNA sequences for seven secretory proteins (reviewed in Pustell *et al.*, 1984; Wieslander *et al.*, 1984; Grond *et al.*, 1987; Dreesen and Case, 1987; Dignam *et al.*, 1989; and Dignam and Case[2]). From these data and from amino acid analyses (Grossbach, 1969; Rydlander, 1984), it is known that the proteins have distinctive amino acid content and arrangement. The spl family (splA, splB, splC, and splD) contains the largest of the proteins, estimated to be from 750 to 1000 kDa in size (Edström, 1980; Hertner *et al.*, 1980; Kao and Case, 1985). Their primary structure is characteristic of fibrous proteins, with periodicities in the amino acid sequence and with a large percentage of charged residues. They consist mainly of segments of 60–90 amino acids, tandemly repeated 75–100 times. Each segment has two domains: a constant (C) and a subrepeat (SR) domain. The subrepeat domain consists of internally repeated segments which tend to have the tripeptide motif (basic residue-proline-acidic residue). The basic residues are typically arginine or lysine, and the acidic residues are glutamate or phosphorylated serine or threonine. The proteins are rich in cysteine; each core repeat of the proteins has 4 invariant cysteines, which are restricted to the C region.

Two other SPs also appear to consist of tandemly repeated segments, although the amino acid content, structure, and size of these repeats are different from those of the spls. sp195[3] has a high content of charged residues, arranged as simple repeats of 25 amino acids (Dreesen *et al.*, 1985; Dreesen and Case, 1987). Each repeat contains 1 cysteine and 1 copy of the tripeptide motif found in the SR domains of spls. sp140 is rich in glycine and charged residues, which are arranged in repeats of 14 amino acids (Dignam *et al.*, 1989). Cysteine is absent, but the tripeptide motif is present in these repeats. The organization of sp185 is different; no sequence repeats are present, but 3 cysteine residues occur approximately every 22 amino acids (Dignam and Case[2]). These data indicate that assembly of SPs could be mediated through numerous intermolecular electrostatic interactions and disulfide bonds.

The functions of the individual proteins, their tertiary structures, and the mechanism of their assembly into silken threads are not known. We initiated studies of *in vitro* assembly, theorizing that intrinsic properties of the proteins directing their *in vitro* behavior could be responsible for directing assembly *in situ*. We used turbidity measurements to investigate assembly and confirmed the results with electron microscopy. From electron micrographs we were able to determine some morphological characteristics of the assembled protein. We also fractionated the proteins so that we might begin to study the roles of individual proteins in the assembly process.

* This research was supported by Contract N00014-87-K-0387 from the Office of Naval Research. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of National Institutes of Health National Research Service Award GM11944.

§ To whom all correspondence should be addressed.

[1] The abbreviations used are: SPs, secretory proteins; spl family, a family of approximately 1000-kDa SPs; C domain, constant domain of an spl; SR domain, subrepeat domain of an spl; MES, 2-(*N*-morpholino)ethanesulfonic acid.

[2] S. Dignam and S. T. Case, submitted for publication.

[3] Our most recent results suggest that previous estimates of the apparent M_r secretory proteins sp130 through sp180 were low. To avoid confusion in comparisons to our previous work, we point out here that sp195 replaces sp180, sp185 replaces sp170, and sp140 replaces sp130.

EXPERIMENTAL PROCEDURES

Larvae—Larvae were reared in our laboratory, as described by Case and Daneholt (1978).

Isolation of Secretory Proteins—Larvae were dissected in an isolation buffer, MKEN, containing 10 mM MES, 100 mM NaCl, 5 mM KCl, 1 mM EDTA at pH 6.3. This isolation buffer differs only slightly from NKE (Mahr *et al.*, 1977); MES was substituted for Tris-maleate (to eliminate UV absorption due to maleate), and the concentration of EDTA was increased to 1 mM. The salivary glands were dissected out, placed in ice-cold MKEN in a siliconized glass depression slide on a cold stage at 0–4 °C, and punctured with a dissecting needle. The glands were stirred with the dissecting needle for 2–3 min to allow the release of lumenal contents. The entire mixture was transferred to a microcentrifuge tube and spun for 30–60 s in a microcentrifuge at 12,000 × *g* to pellet glandular debris. The supernatant, containing secretory proteins, was collected.

Measurement of Turbidity—Turbidity of solutions of secretory proteins was measured in a Cary 219 dual-beam spectrophotometer or a Hewlett-Packard 8452 diode array spectrophotometer.

Measurement of Protein Concentration—The bicinchoninic acid protein assay (Pierce Chemical Co.) was used against a standard of bovine γ-globulin and was calibrated with amino acid analysis. Amino acid analysis was in a Beckman 119CL amino acid analyzer.

Electron Microscopy—Solutions of secretory proteins, usually 2 μg/ml, were placed on squares of Parafilm in 25-μl droplets, covered, and allowed to sit for 2 h. 200-mesh copper grids, with carbon-coated 1% Parlodian support films, were floated on the protein solution for 30 s, rinsed briefly in 90% ethanol, stained in uranyl acetate, 42 μg/ml in 90% ethanol, and rinsed again. Grids were shadowed unidirectionally at an angle of 20° with platinum/palladium, 80:20. Samples were viewed and photographed at 60 kV in a Zeiss EM10 electron microscope. A carbon replica grating with 2160 lines/mm was used to calibrate the magnification of photographed images.

Glycerol Gradient Fractionation—Secretory proteins were fractionated on denaturing gradients of 5–20% glycerol, 5 M guanidine hydrochloride, 100 mM β-mercaptoethanol, 10 mM MES, pH 6.3, 20 mM EDTA. Centrifugation was at 45,000 rpm in an SW60Ti rotor at 4 °C for 43 h. Proteins were prepared for fractionation by adding to the protein solutions described above, crystalline guanidine hydrochloride to a final concentration of 6 M and dithiothreitol to a final concentration of 100 mM. Combined fractions, usually 300–500 μl, were dialyzed against 1 liter of MKEN, pH 6.3, at 4 °C for 14–18 h.

Gel Electrophoresis—Proteins for electrophoresis were precipitated from gradient fractions with 9 volumes of 80% acetone. Protein pellets were washed with 70% acetone, dissolved in sample buffer, boiled for 3 min, and run on 3–20% polyacrylamide gels as described (Kao and Case, 1985). The procedure for casting the gels was modified for the Bio-Rad Mini Protean Cell; 3.65 ml of the 3% acrylamide solution were drawn up into a 5-ml pipette, the pipette was placed into the 20% acrylamide solution, and 0.9 ml was drawn up into the same pipette. The solutions were mixed slightly when two bubbles were allowed to rise through the solutions in the pipette. The solutions were then allowed to run between the assembled glass plates. Gels were run at a constant current of 20 mA each for 25–30 min, fixed in 10% trichloroacetic acid, 10% acetic acid, 30% methanol, and stained with Coomassie Brilliant Blue R.

Circular Dichroic Spectroscopy—Proteins for CD measurements were dialyzed into 10 mM phosphate buffer, pH 6.3, with 90 mM NaF. Measurements were made at room temperature on a Jasco J-500 spectropolarimeter in a cuvette with a path length of 0.1 cm. The fractions of various secondary structures were estimated using the secondary structure estimation programs from Jasco, which are based on computer programs described in Yang *et al.* (1986). The spectra of seven standard proteins were used to calculate the basis spectra: myoglobin, lysozymes, ribonuclease A, papain, cytochrome *c*, hemoglobin, and chymotrypsin C.

RESULTS

SPs Are Isolated as Complexes—The turbidity of solutions containing secretory proteins in MKEN was used to measure assembly of the proteins. Turbidity, the measurement of total light scattered, has been extensively used to monitor assembly of macromolecules including microtubules, actin, and collagen (Andreau and Timasheff, 1980). For assembled proteins that have a long thin rodlike structure, turbidity is predicted to have an inverse third power dependence on wavelength and may be a quantitative measure of the mass of material polymerized (Berne, 1974). Turbidity is typically measured at 350 nm; the sensitivity of the measurement is greater at this than at longer wavelengths, and the interference from absorption in the UV region is avoided.

A wavelength-dependent scan of a solution of SPs is shown in Fig. 1*A*. Absorbance occurs in the UV region, as expected, due to the peptide backbone and amino acid side chains. Phenylalanine, a conserved amino acid in the C domain of spls, may contribute to the absorbance near 260 nm. Absorbance by phenylalanine is usually obscured by absorbance of tyrosine and tryptophan; however, these proteins have little or no tyrosine or tryptophan (Grossbach, 1969). As a consequence, no absorbance peak is seen at 280 nm. A peak of absorbance occurs at 412 nm. This peak coincides with the absorbance maximum of a chromophore in larval hemolymph (data not shown) and does not interfere with measurements of turbidity at 350 nm. A double logarithmic plot of absorb-

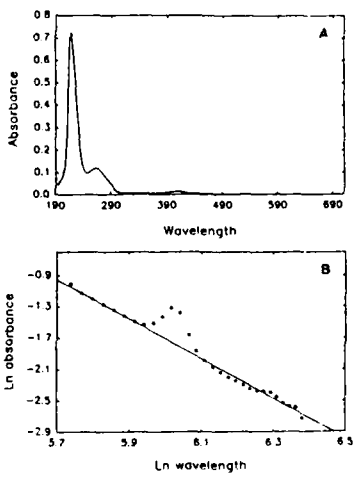

FIG. 1. **Wavelength-dependent scan of SPs.** SPs were prepared as described under "Experimental Procedures" and scanned in a Hewlett-Packard 8452 diode-array spectrophotometer. *A*, *solid line*, scan of SPs at a concentration of 65 μg/ml over the wavelength range 190–700 nm; *dotted line*, scan of SPs at a concentration of 1.03 mg/ml over the wavelength range 300–700 nm, showing the peak of absorbance at 412 nm. *B*, double logarithmic plot of data from *panel A represented by a dotted line* (from 320 to 600 nm). Deviation from a straight line fit to the data occurs at around 412 nm. The slope of the line is approximately −2.5.

ance *versus* wavelength was fit to a straight line with a slope of approximately −2.5 (Fig. 1*B*). Rayleigh scattering of light by small particles should vary as the inverse fourth power of the wavelength; the deviation of SPs from this behavior indicated that SPs might be assembled. A slope of −3 is predicted for proteins with the geometry of long thin rods (Berne, 1974). The geometry of SPs is therefore not that of a long thin rod and could not be determined from this information.

We used electron microscopy to confirm that the observed solution turbidity was due to the presence of assembled structures. Electron micrographs revealed that the proteins isolated at room temperature in MKEN existed as or assembled into complexes (Fig. 2). The most common structure seen in these preparations was a dense network of strands with branched junctions (Fig. 2*A*). Thin smooth fibrils, 5–20 nm in a diameter, and multistanded fibers, 25–200 nm in diameter, usually with interspersed beads of 20–55 nm in diameter, were also seen about once per grid square (Fig. 2, *A* and *B*). These fibrils and fibers also had branched junctions and were continuous with the dense networks. The observed complexes were stable structures. No change in turbidity or appearance in electron micrographs was observed over a period of several days if the protein solutions were prepared and kept at room temperature. No protein was pelleted by centrifugation in a microcentrifuge, 12,000 × *g*, for 15 min, indicating that any assembled complex had a sedimentation coefficient of less than 1,000 S (van de Sande and Jovin, 1982).

To determine if assembly could be slowed to allow the assembly process to be monitored, solutions were prepared at 4 °C and allowed to warm to room temperature. A slight increase in turbidity was observed as the solution warmed up, attaining a maximal value of 20–30% over the original within 2 h (data not shown). However, the turbidity did not decrease when the samples were cooled to 4 °C, indicating that any assembly that had occurred as the solution warmed up was not cold-labile. We could detect no difference between the structures visible in electron micrographs of the proteins before and after the small increase in turbidity that occurred when the solutions were warmed up (data not shown). This increase in turbidity may have indicated that the proteins had assembled further, or it could have resulted from a change in shape or orientation of the macromolecules, which was not detected in electron micrographs.

Disassembly and Reassembly of SP Complexes—Since the electron micrographs confirm that SP complexes are not long thin rods (Fig. 2), turbidity is not, in this instance, a quantitative measure of assembly (Berne, 1974). However, turbidity could be used as a qualitative indication of assembly; for example, turbidity was used to examine the effects of reduction and denaturation on the complexes. The turbidity was reduced to about 75% of the initial values by treatment of the proteins with dithiothreitol. Concomitantly, the complexes that were observed in electron micrographs were significantly altered (Fig. 3*A*). The extensive networks of strands were reduced to small patches of strands with branched junctions and beads or nodules. These patches were often in the shape of circles, which were heterogeneous in size. Treatment with guanidine hydrochloride or with guanidine hydrochloride and dithiothreitol decreased turbidity to about 10% of the initial values and abolished the complexes. In the presence of guanidine hydrochloride only, short strands or fibrils appeared to be attached to spheres 30–45 min in diameter (Fig. 3*B*), while when dithiothreitol was also present, only spheres 35–70 nm in diameter were visible (Fig. 3*C*). After the solutions were dialyzed to remove these reagents, the turbidity, the networks of strands (data not shown), and the complexes of fibrils and beaded fibers were restored (Fig. 3*D*).

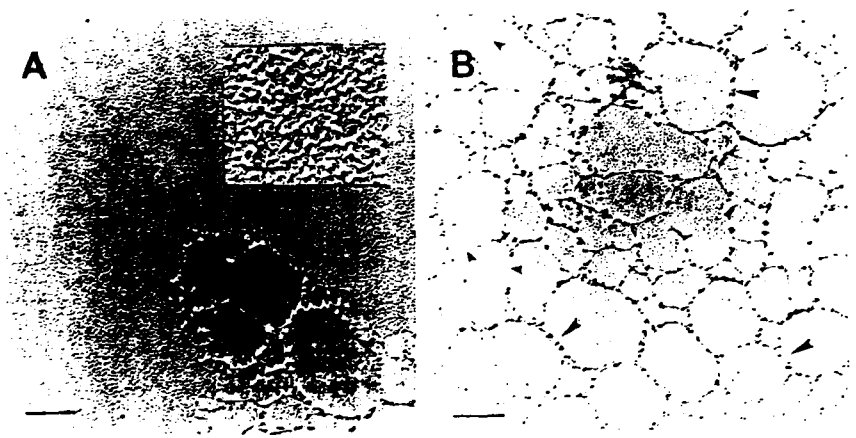

Fig. 2. Electron micrographs of SPs. Solutions of SPs were prepared as described under "Experimental Procedures" and adjusted to a final concentration of 2 μg/ml. *A*, dense networks of strands that were seen most frequently. A portion of the network exhibits patches of fibrils and beaded fibers that were seen, on the average, once per grid square (*bar* = 500 nm). *Inset*, networks at higher magnification (*bar* = 200 nm). *B*, a more extensive area of 25–200-nm diameter fibers with 20–55-nm diameter beads (*large arrowheads*) and 5–20 nm diameter fibrils (*small arrowheads*) that were continuous with networks (*bar* = 500 nm).

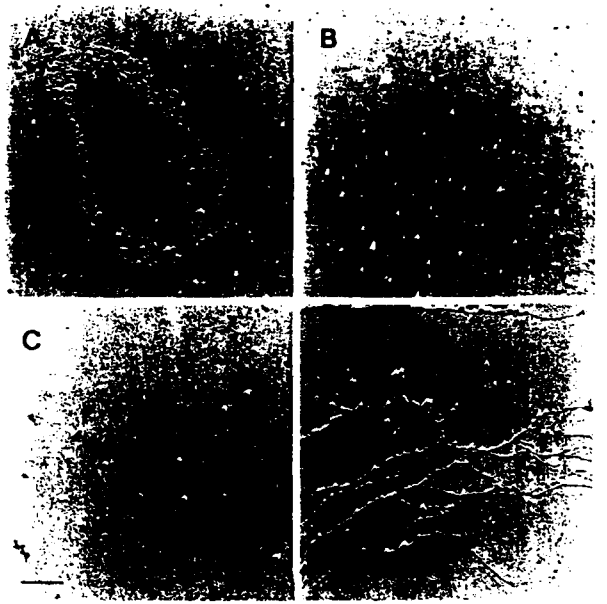

Fig. 3. Electron micrographs of dissociated SP complexes. SPs in MKEN containing 50 mM dithiothreitol (*A*); 5.4 M guanidine hydrochloride (*B*); both 5.4 M guanidine hydrochloride and 50 mM dithiothreitol (*C*). *D*, SPs in MKEN, pH 6.3, after removal of denaturant and reductant by dialysis. *Bar* for all *panels* = 500 nm.

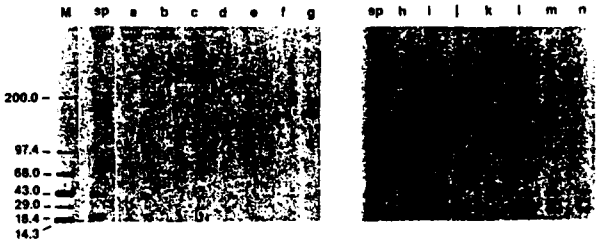

Fig. 4. Fractionation of SPs on denaturing-reducing glycerol gradients; sodium dodecyl sulfate-polyacrylamide gel of gradient fractions. *sp*, unfractionated SPs; *M*, molecular weight markers (Bethesda Research Laboratories). *Numbers* to the *left* are sizes in kDa; *lanes a–n*, gradient fractions, from *bottom* to *top*.

We concluded that the SPs exist in aggregates with regular structure as quickly as we can isolate them in MKEN at pH 6.3. We were unable to block this level of assembly by isolating the proteins at 4 °C nor could we promote disassembly by cooling the protein solutions. These regular structures could be disrupted by denaturation or by denaturation and reduction and were significantly affected by reduction alone. However, solution turbidity and morphologically similar complexes could be restored upon removal of these reagents.

spIs Alone Can Form Complexes—The contribution of individual proteins to the observed structures was investigated. The SPs were denatured, reduced, and fractionated on denaturing glycerol gradients (Fig. 4). The proteins were separated into three size classes: the largest containing only the spIs (at the top of *lanes a–c*), a group containing predominantly sp140 through sp195 (in the middle of *lanes f–i*), and a group containing the smallest proteins, sp15 through sp40 (at the bottom of *lanes l–n*). When proteins from each size class were dialyzed into MKEN, pH 6.3, and viewed by electron microscopy, only pure spIs showed networks of fibrous structures very similar to the structures seen in the total mixture (Fig. 5, *A* and *B*). The dense networks were present; thin smooth fibrils and multistranded beaded fibers with branched junctions were also present. The spI fibers appeared to have nodules rather than distinct beads. The spIs were also capable of forming bundles of fibers in parallel arrays that were 150–200 nm in diameter (Fig. 5*D*). These bundles had only been observed occasionally in unfractionated SPs (Fig. 5*C*). Micrographs of any of the other proteins revealed only small spheres, about 10–30 nm in diameter (Fig. 5, *E* and *F*). From these data, we inferred that the spIs alone are capable of forming the networks, fibrils, and fibers observed in complexes assembled from unfractionated SPs. Furthermore, since the proteins were denatured and reduced in the frac-

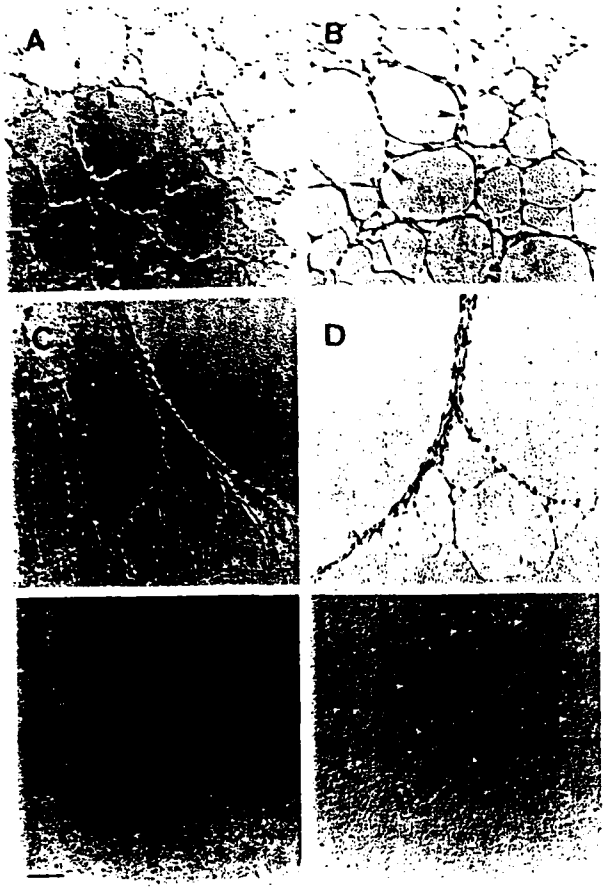

FIG. 5. **Electron micrographs of fractionated and unfractionated SPs.** Unfractionated SPs and SPs fractionated on denaturing gradients (Fig. 4) were dialyzed into MKEN and prepared for electron microscopy. *A*, unfractionated SPs; *B*, purified spIs (*lanes a–c* in Fig. 4). *Large arrowheads* indicate 25–200-nm diameter multistranded beaded fibers, and *small arrowheads* indicate 5–20-nm diameter smooth fibrils. *C*, unfractionated SPs, showing a bundle of fibers; *D*, purified spIs, showing a bundle of fibers; *E*, a mixture of SPs containing predominantly sp140–sp135 (*lanes f–i* in Fig. 4); *F*, a mixture of SPs containing predominantly sp15–sp40 (*lanes l–n* in Fig. 4). *Bar* = 500 nm (*A*, *B*, *C*, *D*) or 200 nm (*E*, *F*).

tionation procedure, the proteins can clearly reassemble.

Secondary Structure of SPs and spIs—Secondary structure of the SP complexes was investigated using CD spectrometry. Spectra of both the unfractionated SPs and the purified reassembled spIs are shown in Fig. 6. The spectra of the total SPs and of the purified spIs had nearly identical features: a positive band at about 190 nm, a prominent negative band near 208 nm, and a negative band appearing as a shoulder centered at 220 nm, with the crossover from positive to negative at 195 nm. These spectra are characteristic of proteins in the $\alpha + \beta$ class of tertiary structure (Manavalan and Johnson, 1983), which are proteins containing distinct regions of α-helix and β-sheet. CD spectra of SPs in the other two size classes did not have characteristic features of this or other known types of regular tertiary structure (data not shown).

The fraction of each of four types of secondary structure in spIs was estimated by fitting the CD spectrum in Fig. 6*B* to reference spectra, using computer programs described by Yang *et al.* (1986). The results indicated that spIs were approximately 15% α-helix, 28–30% β-sheet, 26–28% β-turn, and 25% other secondary structure.

DISCUSSION

From these initial studies of the assembly of secretory proteins, we concluded that the SPs are probably already assembled into complexes in the lumen of salivary glands. These complexes consist of networks of strands, fibrils, and thick fibers, which can be dissociated and reassembled into morphologically similar structures *in vitro*. These web like networks and thickened fibers are clearly not the assembled threads that are spun by larvae but probably represent earlier steps in assembly. These early steps of assembly may occur within the lumen of the salivary glands or even within the secretory cells before secretion into the lumen. In this respect the SPs may resemble collagen, which is assembled from individual chains into a triple helix before it is secreted into that extracellular matrix, at which point it undergoes further assembly. Further assembly of SPs clearly occurs *in vivo*; the ultimate result is the formation of spun threads that are insoluble. We have not yet found conditions that induce the further ordered assembly of SPs *in vitro*.

We also concluded that the spIs constitute the fibrous backbone of the threads. Their abundance, large size, and constitutive expression during larval development had made them favorite candidates for this role, although this had never been shown. The functions of the smaller proteins remain unknown; however, the more recently discovered developmental changes in the expression of some of these proteins (Dreesen and Case, 1987; Dignam *et al.*, 1989) have implicated them in changes in the function and structure of the tubes that occur in larval development.

The CD spectra of SP complexes suggest a tertiary structure with discrete regions of α-helix and β-sheet. This structure is most likely a result of interactions between spI components, which have qualitatively similar spectra. This suggestion is consistent with the secondary structure predictions made by Hamodrakas and Kafatos (1984) in their analysis of the amino acid sequence of spIs. From their analysis, they predicted that each repeat within an spI has a secondary structure of α-helix in the C domain alternating with a regular structure in the SR domain which could not be identified with predictive methods. Their results indicate that the α-helix predicted in the C domain is made up of 18–20% of the amino acids in the entire repeat and that the β-turns predicted in the C domain involve only about 7% of the amino acids in the entire repeat. From comparison of these results to our estimates of 15% α-helix, 28–30% β-sheet, and 26–28% β-turn, obtained from analysis of the CD spectrum, we would predict that: 1) residues in the SR domain do not form α-helices; all or most of the α-helical structure is formed by residues in the C domain; 2) the SR domain has large fractions of β-turn and β-sheet, as few residues in the C domain are predicted to the β-turn and none are predicted to be β-sheet.

Further details of the conformation of spIs and of the assembled threads remain obscure. The exact regions of spIs that interact intermolecularly and intramolecularly, the nature of the interactions, and the conditions that induce further assembly remain to be elucidated. Investigation of these and other questions will be possible with purified spIs.

Acknowledgments—We thank Dr. Jack Correia for advice on the turbidity measurements and for critical comments on the manuscript, Dr. Correia and Dr. Brad Chaires for help in obtaining and interpreting CD data, Dr. Mark Olson for help with amino acid analysis, and Dr. Bill Detrich for initially suggesting that turbidity be used as an assay for assembly.

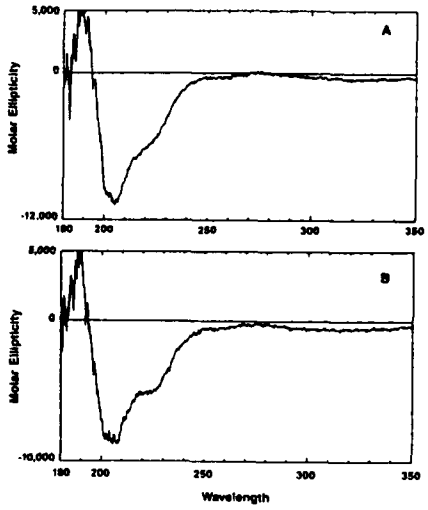

FIG. 6. **CD spectra of SPs.** SPs were dialyzed into 10 mM phosphate buffer, pH 6.3, 100 mM NaF. *A*, spectrum of unfractionated SPs; *B*, spectrum of purified spIs.

REFERENCES

Anderson, S. O. (1970) *Comp. Biochem. Physiol.* **35**, 705–711
Andreau, J. M., and Timasheff, T. N. (1986) *Methods Enzymol.* **130**, 47–59
Berne, B. J. (1974) *J. Mol. Biol.* **89**, 755–758
Case, S. T., and Daneholt, B. (1978) *J. Mol. Biol.* **124**, 221–241
Dignam, S., Yang, L., Lezzi, M., and Case, S. T. (1989) *J. Biol. Chem.* **264**, 9444–9452
Dreesen, T. D., and Case, S. T. (1987) *Gene (Amst.)* **55**, 53–65
Dreesen, T. D., Bower, J. R., and Case, S. T. (1985) *J. Biol. Chem.* **260**, 11824–11830
Dreesen, T. D., Lezzi, M., and Case, S. T. (1988) *J. Cell Biol.* **106**, 21–27
Edstrom, J.-E., Rydlander, L., and Franke, C. (1980) *Chromosoma (Berl.)* **81**, 115–124
Grond, C., Saiga, H., and Edstrom, J. E. (1987) in *Results and Problems in Cell Differentiation* (Hennig, W., ed) Vol. 14, pp. 69–79, Springer-Verlag, Heidelberg, Federal Republic of Germany
Grossbach, U. (1969) *Chromosoma (Berl.)* **28**, 136–187
Grossbach, U. (1977) in *Results and Problems in Cell Differentiation* (Beerman, W., ed) Vol. 8, pp. 147–196, Springer-Verlag, Heidelberg, Federal Republic of Germany
Hamodrakas, S. J., and Kafatos, F. C. (1984) *J. Mol. Evol.* **20**, 296–303
Hertner, T., Meyer, B., Eppenberger, H. M., and Mahr, R. (1984) *Wilhelm Roux's Arch. Dev. Biol.* **189**, 69–72
Kao, W. Y., and Case, S. T. (1985) *J. Cell Biol.* **101**, 1044–1051
Lucas, F., Shaw, J. T. B., and Smith, S. G. (1960) *J. Mol. Biol.* **2**, 339–349
Mahr, R., Lezzi, M., and Eppenberger, H. M. (1977) *J. Cell Sci.* **27**, 1–12
Manavalan, P., and Johnson, W. C. (1983) *Nature* **305**, 831–832
Pustell, J., Kafatos, F. C., Wobus, U., and Baumlein, H. (1984) *J. Mol. Evol.* **20**, 281–295
Rydlander, L. (1984) *Biochem. J.* **220**, 423–431
van de Sande, J. H., and Jovin, T. M. (1982) *EMBO J.* **1**, 115–120
Wallace, J. B., and Merritt, R. W. (1980) *Annu. Rev. Entomol.* **25**, 103–132
Wieslander, L., Hoog, C., Hoog, J.-O., Jornvall, H., Lendahl, U., and Daneholt, B. (1984) *J. Mol. Evol.* **20**, 304–312
Yang, J. T., Wu, C. S.C., and Martinez, H. M. (1986) *Methods Enzymol.* **130**, 208–269